AFRL-VS-TR-1999-1529
E.R.P., No. 1222

A THUNDERSTORM PREDICTION TECHNIQUE BASED ON A PERFECT-PROG APPROACH

Allan J. Bussey
Michael F. Remeika, Major, USAF
Brian Newton, Captain, USAF
Daniel DeBenedictis
Donald C. Norquist

February 1999

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
29 Randolph Rd.
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010

"This technical report has been reviewed and is approved for publication."

DONALD A. CHISHOLM
Chief, Tactical Environmental Support Branch
Battlespace Environment Division

DAVID A. HARDY
Chief, Battlespace Environment Division

This document has been reviewed by the ESC Public Affairs Office (PA) and is releasable to the National Information Service (NTIS).

Qualified requesters may obtain additional copies from the Defense Technical Information Center (DTIC). All others should apply to the NTIS.

If your address has changed, if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify AFRL/VSR, 29 Randolph Road, Hanscom AFB, MA 01731-3010. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	9 Oct 1998	Scientific, Interim, Oct 95-Sep 98

4. TITLE AND SUBTITLE
A Thunderstorm Prediction Technique Based on a Perfect-Prog Approach

5. FUNDING NUMBERS
PE 63707F
PR 4026
TA GT
WU 04

6. AUTHOR(S)
Allan J. Bussey, Michael F. Remeika, Daniel DeBenedictis, Brian Newton, Donald C.Norquist

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/VSBE
29 Randolph Rd
Hanscom AFB, MA 01731-3010

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL-VS-TR-1999-1529
E.R.P., No. 1222

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
We have developed a thunderstorm prediction technique for use with a mesoscale model to satisfy Air Force Weather's Theater Battle Management (TBM) stated requirement for maximum cloud tops and coverage. The prediction technique is to be implemented in the Global Theater/Weather Analysis and Prediction system (GTWAPS). A perfect-prog approach was implemented using synoptic and subsynoptic-scale data for diagnosis and short-range probability forecasts. Future work is discussed, including, most importantly, follow-on testing and refinements of the technique.

14. SUBJECT TERMS
Thunderstorm
Perfect-prog approach
Maximum cloud tops

15. NUMBER OF PAGES
118

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

Contents

Illustrations

Tables

Acknowledgment

We appreciate the interest and support of our branch chief, Don Chisholm. We thank Donna Ruble and Audrey Campana with assisting us with the text preparation. Finally, we thank our colleagues in the Battlespace Environment Division for comments and suggestions.

A Thunderstorm Prediction Technique Based on a Perfect-Prog Approach

1. INTRODUCTION

A thunderstorm prediction technique has been developed using synoptic and subsynoptic-scale data for diagnosis and short-range probability forecasts of any storm (nonsevere and severe) and severe storm occurrence over the continental United States during May, July, and September of 1992 and 1993. The goal was to adopt a methodology different from the currently operational Model Output Statistics (MOS) approach used by the National Weather Service (NWS)/National Center for Environmental Prediction (Reap, 1994)[1] due to the need to address global requirements where real-time or near real-time NEXRAD data, for example, was not available. Our approach limits the number of predictors considered (in the same vein as Albers, 1987)[2]; however, it does not limit the number to include only commonly-used storm-potential indices such as the K Index, Severe Weather Threat Index (SWEAT), Showalter Index, etc. (Knapp, 1995)[3]. We decided against deriving predictors other than those produced by SHARP (NWS' SKEW-T/Hodograph Analysis and Research Program), due to the time involved in using and analyzing predictors for the 12 CONUS regions and three predictand types.

Prediction equations would be used to satisfy Air Force Weather's Theater Battle Management (TBM) stated requirements for maximum cloud tops and coverage and implemented in the Global Theater/Weather Analysis and Prediction System (GTWAPS). Theater forecast

Received for publication 2 October 1998

requirements extend from 0 to 6, 12, 24, and 36 hours; hemispheric requirements include theater and also 48 and 72 hours. GTWAPS is currently planned to be fully operational after the year 2000. The thunderstorm prediction technique described herein was thus tailored for use with a mesoscale model, but not for a specific model, and so the MOS approach was abandoned early in the project for a perfect-prog approach. Other TBM requirements–for clouds, turbulence, icing, surface visibility, and present weather–were addressed by other researchers, in the same Branch at Air Force Research Laboratory, to maximize collaboration and ensure consistency and agreement among the output of the various developmental algorithms.

Follow-on testing using MM5 modeled forecast data from the surface up to 100 mb for August, 1993, is described later in this report.

A stepwise linear regression technique has been used on the most promising predictors for each of 12 CONUS regions and the overall CONUS. Correlation coefficients were first calculated and analyzed to identify the most promising predictors from a set of 121 predictors. Table 1 identifies these predictors, which were generated from SHARP (Hart and Korotky)[4].

2. PREDICTORS

Extensive as the initial pool of predictors was, it did not contain those dynamic parameters such as vertical velocity, differential vorticity advection, and warm air advection, which would be so useful in determining storm initiation. Disregard of such parameters can lead to overforecasting in regions where the tropopause is frequently unstable but dynamics are weak–the result being thunderstorms do not occur nearly so often as one might infer (McGinley, Runk, and Alleca, 1987)[5]. An overlap of the stability and dynamics is necessary to get the best estimates of storm potential. For this study, however, we used a perfect-prog approach in lieu of forecast data in a MOS approach due to constraints mentioned earlier.

Table 1. List of 121 Predictors Reviewed in Developing a Thunderstorm Prediction Technique.

T	surface temp (deg. C)	FF300	300 mb wind speed (knots)
TD	surface dewpoint (deg. C)	DD300	300 mb wind direction (deg.)
TW	surface wetbulb temp (deg. C)	T200	200 mb temp (deg. C)
TE	surface theta-E (deg. K)	HT200	200 mb height (meters, AGL)
HT	surface height (meters, AGL)	FF200	200 mb wind speed (knots)
FF	surface wind speed (knots)	DD200	200 mb wind direction (deg.)
DD	surface wind direction (deg.)	BP	positive buoyant energy (CAPE) (j/kg)
T850	850 mb temp (deg. C)	BM	negative buoyant energy (j/kg)
TD850	850 mb dewpoint (deg. C)	H20IN	precipitable water (in)
TW850	850 mb wetbulb temp (deg. C)	RI	bulk richardson number
TE850	850 mb theta-E (deg. K)	MLI500	negative lifted index to 500 mb
HT850	850 mb height (meters, AGL)	MLI300	negative lifted index to 300 mb
FF850	850 mb wind speed (knots)	MSHOW	negative showalter index
DD850	850 mb wind direction (deg.)	VERT T	vertical totals
T700	700 mb temp (deg. C)	CROSS	cross totals
TD700	700 mb dewpoint (deg. C)	SWEAT	SWEAT index
TW700	700 mb wetbulb temp (deg. C)	KI	K index
TE700	700 mb theta-E (deg. K)	IN13	(unknown)
HT700	700 mb height (meters, AGL)	CAP	cap strength (deg. C)
FF700	700 mb wind speed (knots)	LAPSE	lapse rate (700-500 mb) (deg. C/km)
DD700	700 mb wind direction (deg.)	MAXT	maximum temp forecast (deg. F)
T500	500 mb temp (deg. C)	DD06	0-6 km AGL mean wind direction (deg.)
TD500	500 mb dewpoint (deg. C)	FF06	0-6 km AGL mean wind speed (knots)
TW500	500 mb wetbulb temp (deg. C)	DDSTRM	storm motion (direction, deg.)
TE500	500 mb theta-E (deg. K)	FFSTRM	storm motion (speed, knots)
HT500	500 mb height (meters, AGL)	HEL02	0-2 km AGL total helicity $(m/sec)^2$
FF500	500 mb wind speed (knots)	HEL03	0-3 km AGL total helicity $(m/sec)^2$
DD500	500 mb wind direction (deg.)	SHR02	0-2 km AGL total shear (10^{-3} sec^{-1})
T300	300 mb temp (deg. C)	SHR03	0-3 km AGL total shear (10^{-3} sec^{-1})
TD300	300 mb dewpoint (deg. C)	SHRBRN	BRN shear $(m/sec)^2$
TW300	300 mb wetbulb temp (deg. C)	EHI	energy/helicity index
TE300	300 mb theta-E (deg. K)	DDSR	0-3 km AGL, SR directional change (deg.)
HT300	300 mb height (meters, AGL)	TEI	theta-E index (deg. C)

Table 1. continued

TEMAXT	max theta-E (deg. K)
TEMAXP	max theta-E pressure (mb)
TEIMINT	min theta-E (deg. K)
TEIMINP	min theta-E pressure (mb)
IN35	(unknown)
PSFC	surface pressure (mb)
PMAX	Wmax pressure (mb)
LCLP	LCL pressure (mb)
WBZP	wetbulb zero pressure (mb)
WBZT	wetbulb zero height (ft, AGL)
MELP	negative equilibrium lvl pres.(mb)
MPLP	maximum parcel level pressure (mb)
ELHT	equilibrium level height (ft. AGL)
MPLHT	max. parcel level height (ft, AGL)
RHEL2	rel. helicity, sfc to 2 km AGL $(m/sec)^2$
SHRPM2	positive shear, sfc to 2 km AGL(10^{-3} sec^{-1})
SHR2	neg. shear, sfc to 2 km AGL (10^{-3} sec^{-1})
HELPM2	pos. helicity, sfc to 2 km AGL $(m/sec)^2$
HELM2	neg. helicity, sfc to 2km AGL $(m/sec)^2$
RHEL4	rel. helicity, sfc to 4 km AGL $(m/sec)^2$
SHRPM4	pos. shear, sfc to 4 km AGL (10^{-3} sec^{-1})
SHRM4	neg. shear, sfc to 4 km AGL (10^{-3} sec^{-1})
HELPM4	pos. helicity, sfc to 4 km AGL $(m/sec)^2$
HELM4	neg. helicity, sfc to 4 km AGL $(m/sec)^2$
RHEL6	rel. helicity, sfc to 6 km AGL $(m/sec)^2$
SHRPM6	pos. shear, sfc to 6 km AGL (10^{-3} sec^{-1})
SHRM6	neg. shear, sfc to 6 km AGL (10^{-3} sec^{-1})
HELPM6	pos. helicity, sfc to 6 km AGL $(m/sec)^2$
HEL6	neg. helicity, sfc to 6 km AGL $(m/sec)^2$
VH500	horiz. vort., sfc to 500 m AGL (10^{-3} sec^{-1})
VS500	strmwise vort, sfc to 500m AGL (10^{-3} sec^{-1})
FF500	mean inflow speed (knots)
DD500	mean inflow direction (deg.)
VH1000	horiz. vort., sfc to 1000 m AGL (10^{-3} sec^{-1})
VS1000	strmwse vort., sfc to 1000 m AGL (10^{-3} sec^{-1})
FF1000	mean inflow speed, sfc to 1000 m AGL (knots)
DD1000	mean inflow direction, sfc to 1000 m AGL (deg.)
VH1500	horiz. vort., sfc to 1500 m AGL (10^{-3} sec^{-1})
VS1500	strmwse vort., sfc to 1500 m AGL (10^{-3} sec^{-1})
FF1500	mean inflow speed, sfc to 1500 m AGL (knots)
DD1500	mean inflow direction, sfc to 1500 m AGL (deg.)
VH2000	horiz. vort., sfc to 2000 m AGL (10^{-3} sec^{-1})
VS2000	strmwse vort., sfc to 2000 m AGL (10^{-3} sec^{-1})
FF2000	mean inflow speed, sfc to 2000 m AGL (knots)
DD2000	mean inflow direction, sfc to 2000 m AGL (deg.)
VH2500	horiz. vort., sfc to 2500 m AGL (10^{-3} sec^{-1})
VS2500	strmwse vort., sfc to 2500 m AGL (10^{-3} sec^{-1})
FF2500	mean inflow speed, sfc to 2000 m AGL (knots)
DD2500	mean inflow direction, sfc to 2500 m AGL (deg.)
VH3000	horiz. vort., sfc to 3000 m AGL (10^{-3} sec^{-1})
VS3000	strmwse vort., sfc to 3000 m AGL (10^{-3} sec^{-1})
FF3000	mean inflow speed, sfc to 3000 m AGL (knots)
DD3000	mean inflow direction, sfc to 2500 m AGL (deg.)

A subset of the parameters in Table 1 was constructed after determining the most promising predictors, and then stepwise regression was performed on the predictor/predictand data set to obtain predictive equations. Before discussing how the predictand was determined and produced, first we discuss how our pool of predictors was generated.

2.1 SHARP Program

The SHARP program was chosen to generate predictors for this study because of its ease of use and plentiful choice of output parameters. As long as upper air sounding data is formatted correctly, output can be generated relatively easily and, if desired, copied to a database file for statistical analysis. Soundings can also be displayed and edited if desired and a generous sample of derived output parameters can be displayed for quick analysis. It is considered an ideal tool for researchers.

The challenge was formatting a large volume of soundings into the desired configuration (see Figure 1). SHARP requires height (m), temperature (deg. C), dewpoint temperature (deg. C), wetbulb temperature (deg. C), wetbulb potential temperature (deg. C), wind direction (deg.) and wind speed (knots), at the surface and interpolated every 50 mb up to 100 mb. SHARP also requires pressure (mb), wind direction (deg.), and wind speed (knots) at the surface and interpolated every 500 m up to 9000 m; and finally storm direction (deg.) and speed (knots). As our RAOB data were not collected from the AFOS system which is the target operational system for which SHARP was designed and implemented, but rather from the database archive at AF Combat Climatology Center (AFCCC), it was necessary to reformat the data from DATSAV2 to SHARP text file format. The files were then moved or copied into the C:\sharp\raobs directory (for Pcs) for processing. SHARP processing methodology is fully described in the SHARP Workstation v1.50 Users Manual. After specifying the number and types of database elements and then databasing (a) sounding(s), the output can be viewed directly in C:\sharp\dbase1.txt or input to an Excel spreadsheet for editing and further processing.

100489001.LBF							
LEVEL	HEIGH	TEMP	DEW F	WETBU	WPT	W DIR	W SPD
SFCLVL							
922.0,	847.0,	8.6,	-5.4,	2.5,	6.8,	215.0,	7.0
%END%							
TROP							
-999.0,	-999.0,	-999.0,	-999.0,	-999.0,	-999.0,	-999.0,	-999.0
%END%							
WPTMAX							
300.0,	9360.0,	-42.9,	-72.9,	-43.4,	16.5,	280.0,	73.1
%END%							
SMOOTHED							
922.0,	847.0,	8.6,	-5.4,	2.5,	6.8,	215.0,	7.0
900.0,	1045.7,	8.2,	-8.0,	1.5,	7.0,	219.3,	11.1
850.0,	1516.0,	3.8,	-8.2,	-1.0,	7.2,	245.0,	15.0
800.0,	2003.0,	0.6,	-10.4,	-3.4,	7.6,	283.6,	25.7
750.0,	2521.3,	0.6,	-12.6,	-4.1,	9.7,	290.0,	34.8
700.0,	3076.0,	-1.3,	-14.3,	-5.8,	11.1,	295.0,	37.9
650.0,	3654.3,	-2.9,	-14.8,	-7.0,	13.1,	304.9,	39.0
600.0,	4278.5,	-7.4,	-17.6,	-10.5,	13.5,	305.2,	46.1
550.0,	4957.1,	-9.8,	-27.8,	-13.8,	14.1,	305.0,	52.5
500.0,	5700.0,	-15.3,	-30.3,	-18.1,	14.3,	305.0,	56.0
450.0,	6474.5,	-21.7,	-35.1,	-23.5,	14.5,	295.9,	66.4
400.0,	7340.0,	-26.9,	-41.9,	-28.3,	15.5,	285.0,	74.0
350.0,	8277.3,	-33.5,	-63.5,	-34.6,	16.1,	282.8,	72.5
300.0,	9360.0,	-42.9,	-72.9,	-43.4,	16.5,	280.0,	73.1
250.0,	10560.0,	-52.3,	-999.0,	-999.0,	-999.0,	280.0,	70.9
200.0,	11970.0,	-61.5,	-999.0,	-999.0,	-999.0,	265.0,	81.0
150.0,	13760.0,	-58.7,	-999.0,	-999.0,	-999.0,	285.0,	73.1
100.0,	16270.0,	-65.7,	-999.0,	-999.0,	-999.0,	280.0,	63.9
%END%							
WINDS							
922.0,	847.0,	-999.0,	-999.0,	-999.0,	-999.0,	215.0,	7.0
905.0,	1000.0,	-999.0,	-999.0,	-999.0,	-999.0,	217.8,	10.4
851.7,	1500.0,	-999.0,	-999.0,	-999.0,	-999.0,	243.9,	14.9
800.3,	2000.0,	-999.0,	-999.0,	-999.0,	-999.0,	283.4,	25.6
752.0,	2500.0,	-999.0,	-999.0,	-999.0,	-999.0,	290.0,	34.6
706.6,	3000.0,	-999.0,	-999.0,	-999.0,	-999.0,	293.9,	37.7
663.0,	3500.0,	-999.0,	-999.0,	-999.0,	-999.0,	302.3,	38.7
621.8,	4000.0,	-999.0,	-999.0,	-999.0,	-999.0,	305.0,	43.0
583.1,	4500.0,	-999.0,	-999.0,	-999.0,	-999.0,	308.8,	46.8
547.0,	5000.0,	-999.0,	-999.0,	-999.0,	-999.0,	305.0,	52.7
513.0,	5500.0,	-999.0,	-999.0,	-999.0,	-999.0,	305.0,	55.0
480.0,	6000.0,	-999.0,	-999.0,	-999.0,	-999.0,	301.2,	59.8
448.4,	6500.0,	-999.0,	-999.0,	-999.0,	-999.0,	295.6,	66.8
419.0,	7000.0,	-999.0,	-999.0,	-999.0,	-999.0,	290.1,	73.9
391.0,	7500.0,	-999.0,	-999.0,	-999.0,	-999.0,	285.0,	72.9
364.1,	8000.0,	-999.0,	-999.0,	-999.0,	-999.0,	283.7,	72.3
339.1,	8500.0,	-999.0,	-999.0,	-999.0,	-999.0,	282.1,	72.6
315.8,	9000.0,	-999.0,	-999.0,	-999.0,	-999.0,	280.5,	73.0
%END%							
STORM							
246,	24						
%END%							

Figure 1. Sample SHARP-Formatted Upper-Air RAOB Observation for LBF, North Platte, NE on 4 Oct 89 at 00Z.

2.2 Upper Air Soundings

Efficient reformatting of DATSAV2 upper air sounding data, for SHARP use, required writing Visual Basic macro code to reformat all upper air data available during the months of May, July, and September 1992 and 1993 for the 71 sites chosen for this study. These sites were selected due to data availability. The 71 station sites were then separated into 12 regions (see Figure 2), in agreement with categorization of the CONUS by a research team doing similar work to predict other present weather and surface visibility. A list of the stations in each region is shown in Table 2.

Decisions were made to use as many soundings as possible while maintaining data integrity (2,130 soundings out of 25,560, or 8.3%, were discarded). There had to be at least 15 groups (representing levels) of data; otherwise, the sounding was not used. If there was no data above 200 mb, the sounding was not used. Likewise, there had to be some data present in the lowest 150 mb. The amount of missing data required us to make other assumptions:

1) Extrapolate data down to surface level to fill in missing surface T, TD. The standard lapse rate was used to help approximate surface temperature. For morning soundings, this ignores potential radiation inversion, which results in a more unstable profile for the few soundings affected. The hydrostatic approximation was used to replace missing surface pressure;

2) Pull down wind direction and speed from one of the next six highest data groups to fill missing surface wind direction and speed;

3) Pull up dewpoint from a lower level to fill in 300 mb TD;

4) Pull up wind direction and speed from one of the next six highest data groups to fill in 100 mb wind direction and speed. Only temperature and dewpoint were extrapolated, by using standard lapse rate.

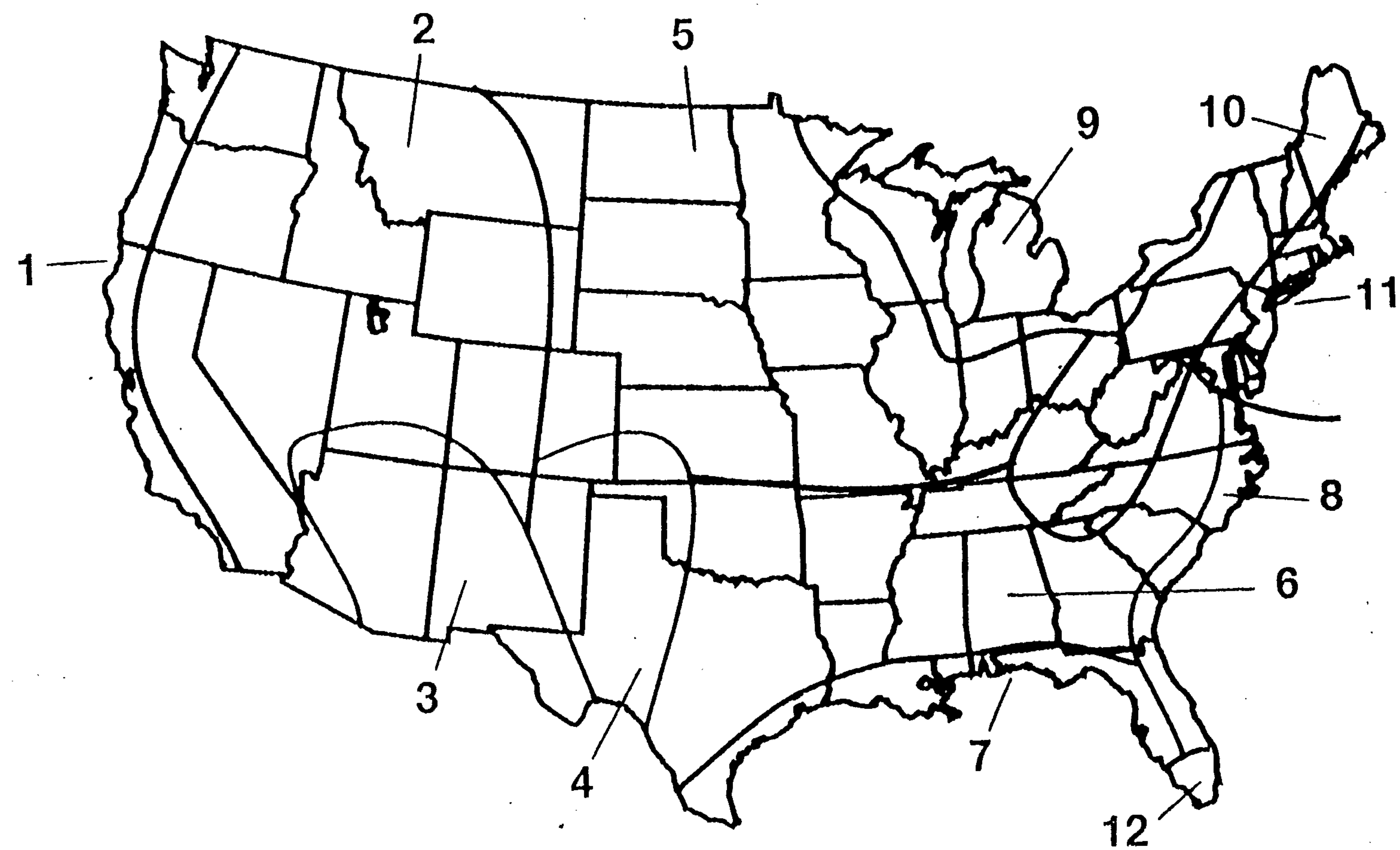


Figure 2. Twelve CONUS Climatological Regions.

Table 2. List of 71 RAOB Station Sites in Each of 12 CONUS Regions

REGION 1 (WEST COAST)						
NTD	723910	PT MUGU (NAWS)	CA	34.11N	119.11W	2 (M) SA,RB
OAK	724930	OAKLAND	CA	37.73N	122.21W	3 (M) SA,RB
SLE	726940	SALEM/MCNARY	OR	44.91N	123.00W	61 (M) SA,RB,SY
UIL	727970	QUILLAYUTE STATE	WA	47.95N	124.55W	62 (M) SA,RB,SY
VBG	723930	VANDENBERG AFB	CA	34.73N	120.58W	112 (M) SA,RB,RR
REGION 2 (INTERMOUNTAIN WEST)						
BOI	726810	BOISE MUNICIPAL	ID	43.56N	116.21W	874 (M) SA,RB,TJ,GU,MO,SY
DRA	723870	MERCURY/DESERT ROCK	NV	36.61N	116.01W	1009 (M) SA,RB
EDW	723810	EDWARDS AFB	CA	34.90N	117.88W	702 (M) SA,RB,GU
ELY	724860	ELY/YELLAND FLD	NV	39.28N	114.85W	1909 (M) SA,RB,SY
GEG	727850	SPOKANE INTL ARPT	WA	47.63N	117.53W	721 (M) SA,RB,TJ,GU,MO,SY
GJT	724760	GRAND JUNCTION	CO	39.11N	108.53W	1475 (M) SA,RB,MO,SY
GTF	727750	GREAT FALLS INTL	MT	47.48N	111.36W	1115 (M) SA,RB,TJ,GU,SY
LND	725760	LANDER/HUNT FIELD	WY	42081N	108.73W	1694 (M) SA,RB,TJ,SY
MFR	725970	MEDFORD/JACKSON CO.	OR	42.36N	122.86W	405 (M) SA,RR,RB,TJ,GU,MO,SY
SLC	725720	SALT LAKE CITY INTL	UT	40.78N	111.96W	1288 (M) SA,RB,TJ,GU,MO,SY
WMC	725830	WINNEMUCCA MUNI	NV	40.90N	117.80W	1315 (M) SA,RB,SY
REGION 3 (SOUTHWEST)						
ABQ	723650	ALBUQUERQUE INTL	NM	35.05N	106.61W	1620 (M) SA,RR,RB,TJ,GU,MO,SY
ELP	722700	EL PASO INTL ARPT	TX	31.80N	106.40W	1194 (M) SA,RB,TJ,GU,SY
FHU	722730	FORT HUACHUCA/LIBBY	AZ	31.60N	110.35W	1438 (M) SA,RB
INS	746140	INDIAN SPRINGS GNRY	NV	36.58N	115.66W	952 (M) SA,RB
TUS	722740	TUCSON INTL ARPT	AZ	32.11N	110.93W	779 (M) SA,RR,RB,MO,SY
REGION 4 (PANHANDLE)						
AMA	723630	AMARILLO ARPT (AWOS)	TX	35.23N	101.70W	1099 (M) SA,RR,RB,MO,SY
DDC	724510	DODGE CITY (AWOS)	KS	37.76N	99.96W	790 (M) SA,RR,RB,TJ,GU,SY
DRT	722610	DEL RIO INTL (AUT)	TX	29.36N	100.91W	307 (M) SA,RB,GU,SY
MAF	722650	MIDLAND REGIONAL	TX	31.95N	102.18W	872 (M) SA,RR,RB,SY
REGION 5 (MIDWEST & NORTHERN TIER)						
BIS	727640	BISMARCK MUNICIPAL	ND	46.76N	100.75W	506 (M) SA,RR,RB,TJ,GU,SY
BNA	723270	NASHVILLE METRO	TN	36.13N	86.68W	180 (M) SA,RR,RB,GU,MO,SY
DAY	724290	DAYTON/JAMES M. COX	OH	39.90N	84.20W	306 (M) SA,RB,GU,MO,SY
DEN	725650	DENVER INTL ARPT	CO	39.86N	104.66W	1656 (M) SA,SY
GGW	727680	GLASGOW INTL ARPT	MT	48.21N	106.61W	700 (M) SA,RB,SY
HON	726540	HURON REGIONAL	SD	44.38N	98.21W	393 (M) SA,RR,RB,SY
LBF	725620	N. PLATTE/LEE BIRD	NE	41.13N	100.68W	849 (M) SA,RR,RB,TJ,GU,SY
OMA	725500	OMAHA/EPPLEY FLD	NE	41.30N	95.90W	299 (M) SA,RR,GU,MO,SY
PAH	724350	PADUCAH/BARKLEY	KY	37.06N	88.76W	126 (M) SA,RR,RB
PIA	725320	PEORIA MUNICIPAL	IL	40.66N	89.68W	202 (M) SA,RB,TJ,SY
RAP	726620	RAPID CITY REGIONAL	SD	44.05N	103.06W	966 (M) SA,RR,RB,TJ,GU,SY
STC	726550	ST CLOUD MUNICIPAL	MN	45.55N	94.06W	312 (M) SA,RB
TOP	724560	TOPEKA/BILLARD MUNI	KS	39.06N	95.61W	270 (M) SA,RR,RB,TJ,GU,SY
UMN	723440	MONET	MO	36.88N	93.90W	437 (M) RR,RB,TJ,SY
REGION 6 (SOUTHWEST INTERIOR)						
CKL	722290	CENTREVILLE/BIBB CO.	AL	32.90N	87.25W	140 (M) SA,RR,RB
GGG	722447	LONGVIEW/GREGG CO.	TX	32.38N	94.71W	111 (M) SA,RR
GSO	723170	GREENSBORO/PIEDMONT	NC	36.08N	79.95W	270 (M) SA,RB,TJ,MO,SY
JAN	722350	JACKSON/THOMPSON	MS	32.21N	90.08W	100 (M) SA,RR,RB,TJ,GU,SY
LZK	723400	NORTH LITTLE ROCK	AR	34.83N	92.25W	165 (M) RR,RB,SY
OUN	723570	NORMAN/WESTHEIMER	OK	35.21N	97.45W	357 (M) RB
SEP	722600	STEPHENVILLE/CLARK	TX	32.21N	98.18W	402 (M) SA,RR,RB
REGION 7 (GULF COAST)						
BRO	722500	BROWNSVILLE INTL	TX	25.90N	97.43W	6 (M) SA,RR,RB,TJ,GU,SY
CRP	722510	CORPUS CHRISTI INTL	TX	27.76N	97.50W	13 (M) SA,RR,RB,SY
LCH	722400	LAKE CHARLES MUNI	LA	30.11N	93.21W	10 (M) SA,RR,RB,SY
LIX	722330	SLIDELL MUNICIPAL	LA	30.33N	89.81W	8 (M) RB
TBW	722100	TAMPA BAY AREA	FL	27.70N	82.40W	13 (M) RR,RB
TLH	722140	TALLAHASSEE RGNL	FL	30.38N	84.36W	21 (M) SA,RB,TJ,GU,SY
REGION 8 (SOUTHEAST COAST)						
AYS	722130	WAYCROSS/WARE CO.	GA	31.25N	82.40W	46 (M) SA,RR,RB
CHS	722080	CHARLESTOWN MUNI	SC	32.90N	80.03W	15 (M) SA,RR,RB,SY
HAT	723040	CAPE HATTERAS	NC	35.26N	75.55W	3 (M) SA,RR,RB,GU,SY
PBI	722030	WEST PALM BEACH	FL	26.68N	80.11W	6 (M) SA,RR,RB,SY
REGION 9 (GREAT LAKES)						
BUF	725280	BUFFALO INTL ARPT	NY	42.93N	78.73W	215 (M) SA,RR,RB,TJ,GU,MO,SY
FNT	726370	FLINT/BISHOP INTL	MI	42.96N	83.75W	233 (M) SA,RB,TJ,MO,SY
GRB	726450	GREEN BAY/STRAUBEL	WI	44.48N	88.13W	214 (M) SA,RB,TJ,SY
INL	727470	INTERNATIONAL FALLS	MN	48.56N	93.38W	361 (M) SA,RB,TJ,GU,SY
SSM	727340	SAULT STE MARIE	MI	46.46N	84.36W	221 (M) SA,RB,TJ,GU,SY
REGION 10 (APPALACHIANS)						
AHN	723110	ATHENS MUNICIPAL	GA	33.95N	83.31W	247 (M) SA,RR,RB,SY
ALB	725180	ALBANY COUNTY ARPT	NY	42.75N	73.80W	89 (M) SA,RR,RB,TJ,GU,SY
CAR	727120	CARIBOU MUNICIPAL	ME	46.86N	68.01W	190 (M) SA,RB,TJ,GU,SY
HTS	724250	HUNTINGTON/TRI STATE	WV	38.36N	82.55W	255 (M) SA,RB,SY
PIT	725200	PITTSBURGH INTL	PA	40.50N	80.21W	373 (M) SA,RR,RB,TJ,GU,MO,SY
REGION 11 (NORTHEAST COAST)						
ACY	702470	ATLANTIC CITY INTL	NJ	39.45N	74.56W	20 (M) SA,RB,RR
IAD	724030	WASHINGTON/DULLES	VA	38.95N	77.45W	98 (M) SA,RB,MO,SY
PWM	726060	PORTLAND INTL	ME	43.65N	70.31W	19 (M) SA,RR,RB,TJ,GU,SY
WAL	724020	WALLOPS ISLAND	VA	37.93N	75.48W	41 (M) SA,RB
REGION 12 (SOUTHEAST)						
EYW	722010	KEY WEST INTL	FL	24.55N	81.75W	6 (M) SA,RR,RB,SY

5) Interpolate between layers to fill in missing T, TD, TW, FF, and DD; and

6) Highest data group needs to be higher than 200 mb. Since SHARP requires data up to and including 100 mb, one wouldn't want to extrapolate across a 5 to 6 km vertical extent in the top of the sounding, much less in the lower part of the sounding.

Once the steps listed above were accomplished, and RAOB-level data filled in, the second interpolation was carried out to provide SHARP-level data for pressure levels shown in the upper half of Figure 1. Interpolation required use of standard logarithms, since data was plotted on Skew-T/log P diagrams.

Producing data in similar format as the bottom part of Figure 1 required interpolation of T, DD, and FF to the desired height levels, which was accomplished in much the same way as the step described above.

Hand-calculated interpolations and manual Skew-T/log P plots of sample RAOB raw sounding data agreed with similar calculations and plots of analogous SHARP-formatted data sets, thus verifying our interpolation algorithm. A comparison of a sample RAOB and SHARP-formatted sounding is shown in Figure 3. The sample RAOB is shown by solid lines, and the SHARP sounding (dashed lines) adds a dewpoint at 300 mb and temperature and dewpoint at the surface for this example, which is the 3 May 92 00Z Rapid City, South Dakota profile. As one might expect, the derived soundings are biased by being too moist at 300 mb and/or at the surface, where our algorithm needed to fill in missing data. However, from having looked at a large sampling of the data, this was the case in only a small minority of the soundings. We conclude that the hydrostatic approximation in calculating surface temperature, together with dropping down TD from a higher level would result in added surface layer instability, thus potentially leading to overforecasting storm occurrence.

Only 00Z and 12Z soundings were used, and any 18Z soundings discarded, producing a maximum of 25,560 soundings (two soundings per day, up to 30 days per month, three months per year, two years, and 71 RAOB stations, minus the unusable RAOBs). The actual number of soundings available after editing and used for statistical analysis was 23,430. By region, the totals are shown in Table 3.

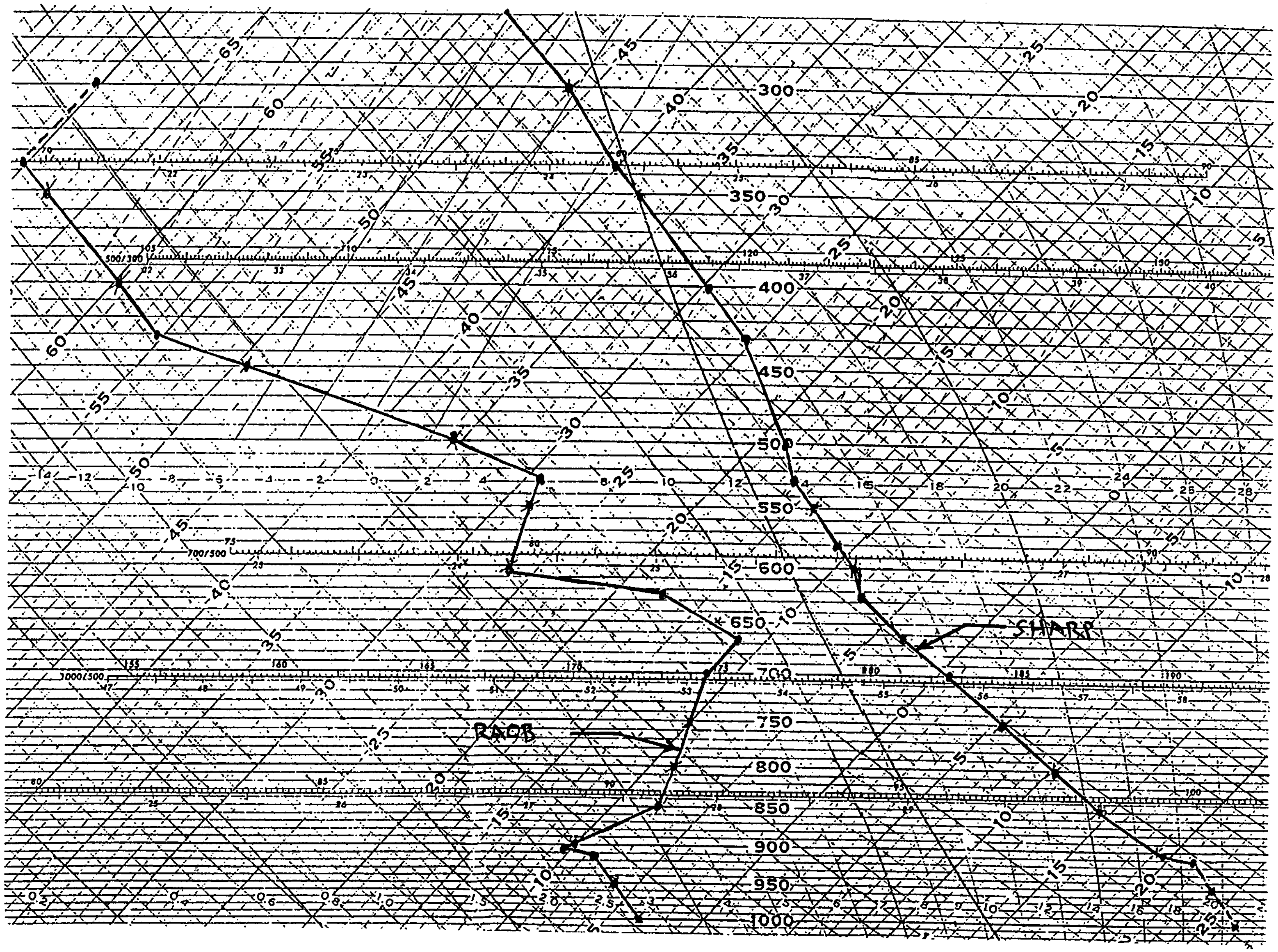


Figure 3. Comparison of Sample RAOB and SHARP-Formatted Sounding.

Table 3. Number of RAOB Soundings Produced for Each of 12 CONUS Regions

	1992 (May, July, Sep.)	1993 (May, July, Sep)	Total	Percent
Region 1	674	681	1355	5.8
Region 2	1812	1783	1395	15.3
Region 3	721	726	1447	6.2
Region 4	648	698	1346	5.8
Region 5	2427	2409	4836	20.6
Region 6	1173	1200	2373	10.1
Region 7	1016	1032	2048	8.7
Region 8	683	701	1384	5.9
Region 9	840	857	1697	7.2
Region 10	841	833	1674	7.1
Region 11	660	679	1339	5.7
Region 12	165	171	336	1.4
Total	11,660	11,770	23,430	99.8

3. PREDICTANDS

Predictands used overall and for each region are illustrated in Figures 4-16.

For these figures, LTG (lightning), OBS (surface station observations), RAD (manually digitized radar), STORM (NCDC storm reports), *OBS (surface station obs of severe storms), *RAD (manually digitized radar obs of severe storms), and *STORM (NCDC severe storm reports) are raw storm predictands, and HITS (storm w/i 33 s.m. and ± 1 hour), sv HITS (same as HITS but for severe storm), HITM (storm w/i 66 s.m. and ± 3 hours), sv HITM (same as HITM but for severe storm), and HITL (storm w/i 100 s.m. and ± 6 hours), sv HITL (same as HITL but for severe storm) are derived predictands. Derived predictands are dependent on any storm occurrence (using LTG, OBS, RAD, STORM) and/or severe storm occurrence (using *OBS, *RAD, *STORM). The -S, -M, and -L suffixes represent a further categorization, described in detail later in this section.

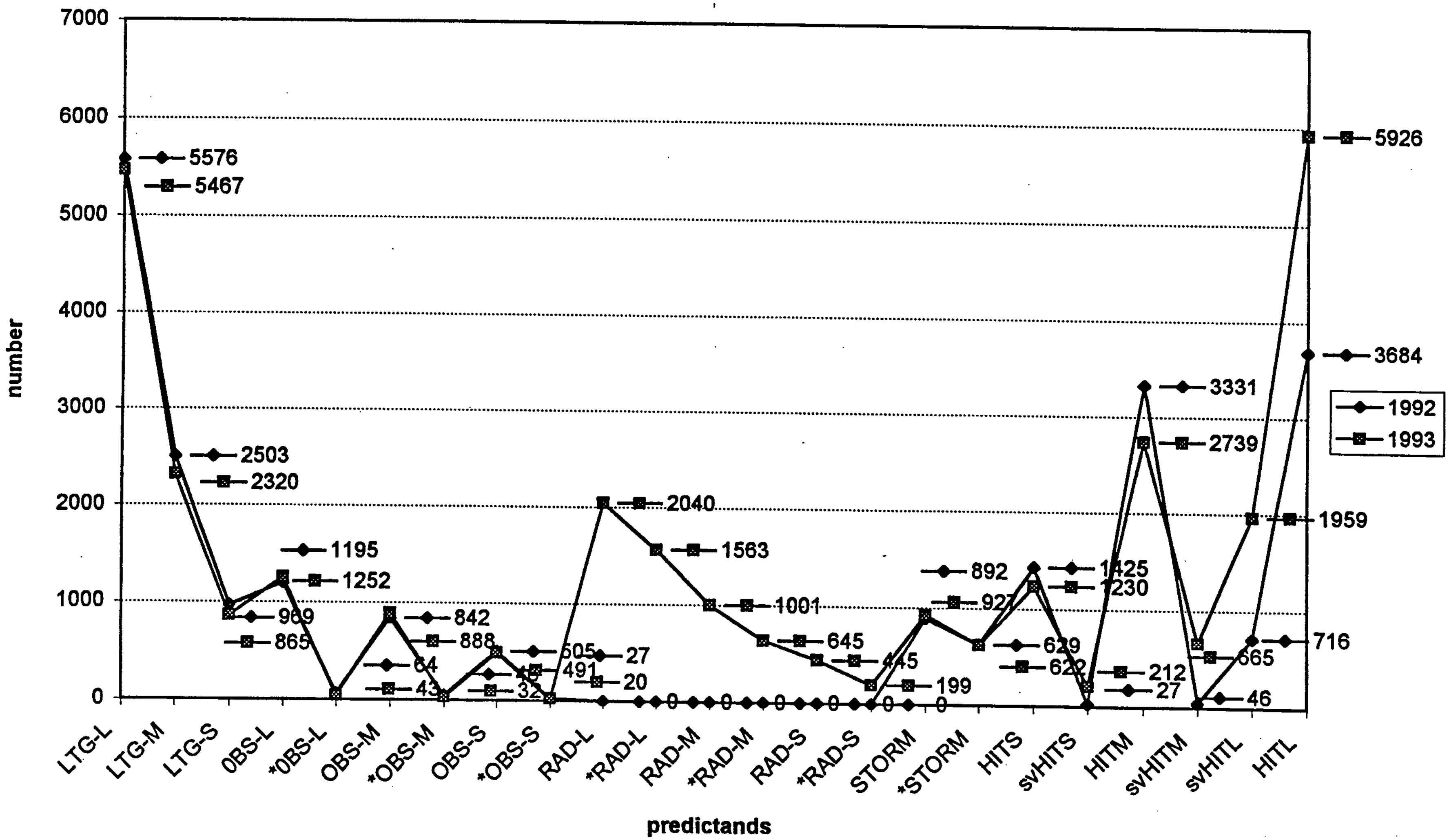


Figure 4. Predictands and Number for Overall Case.

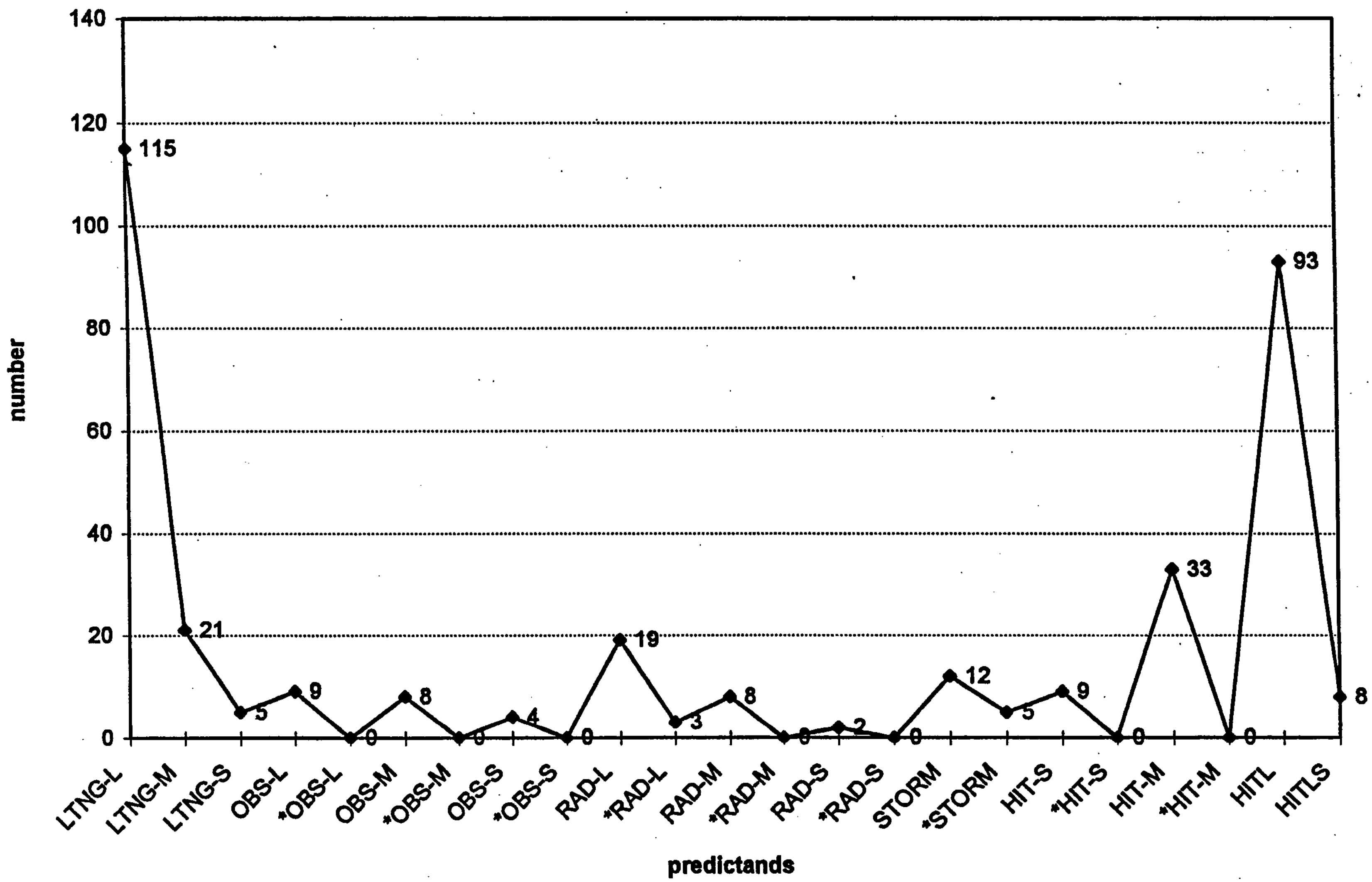


Figure 5. Same as in Figure 4 but for Region 1 (West Coast).

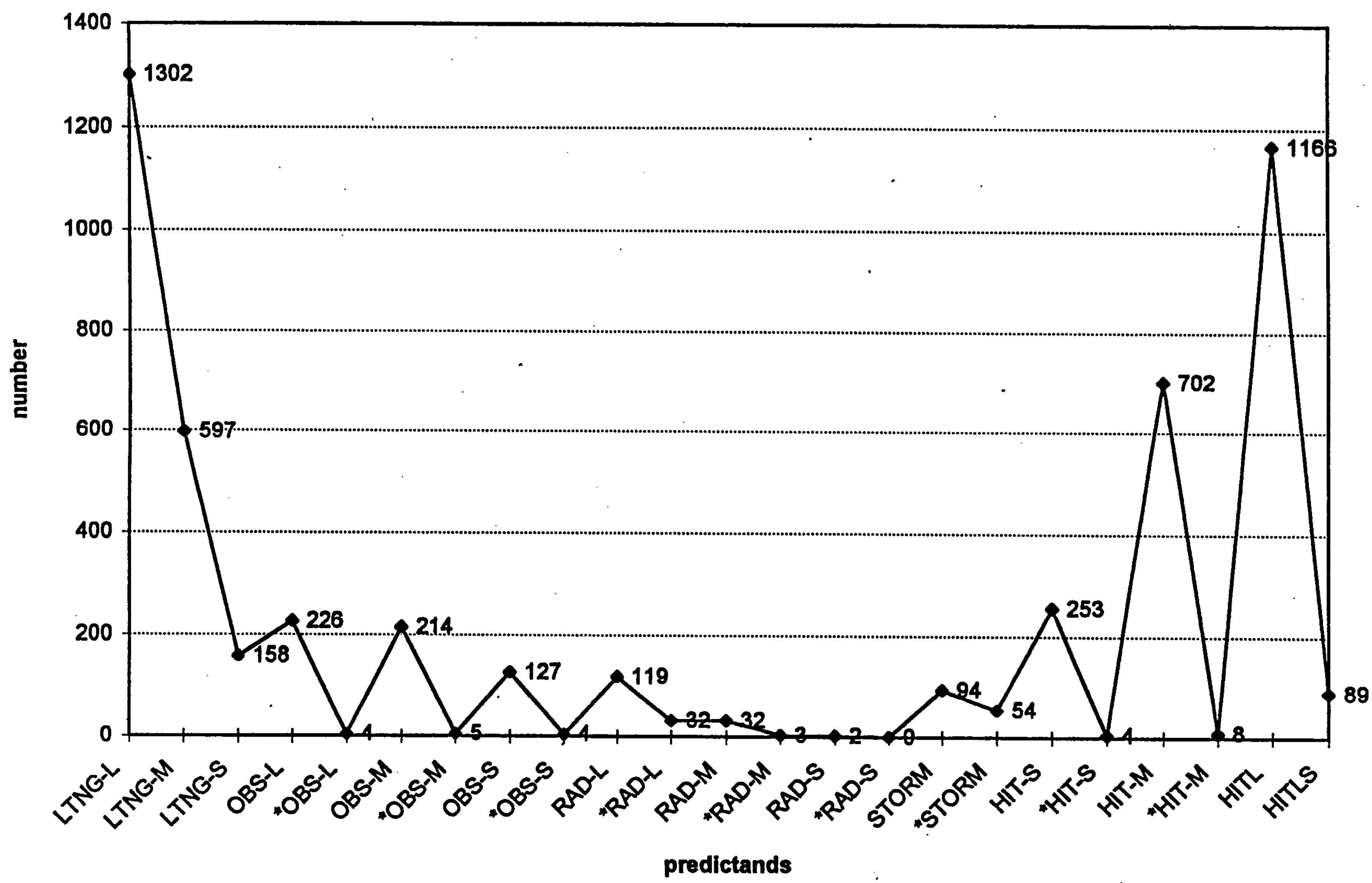


Figure 6. Same as in Figure 4 but for Region 2 (Intermountain West).

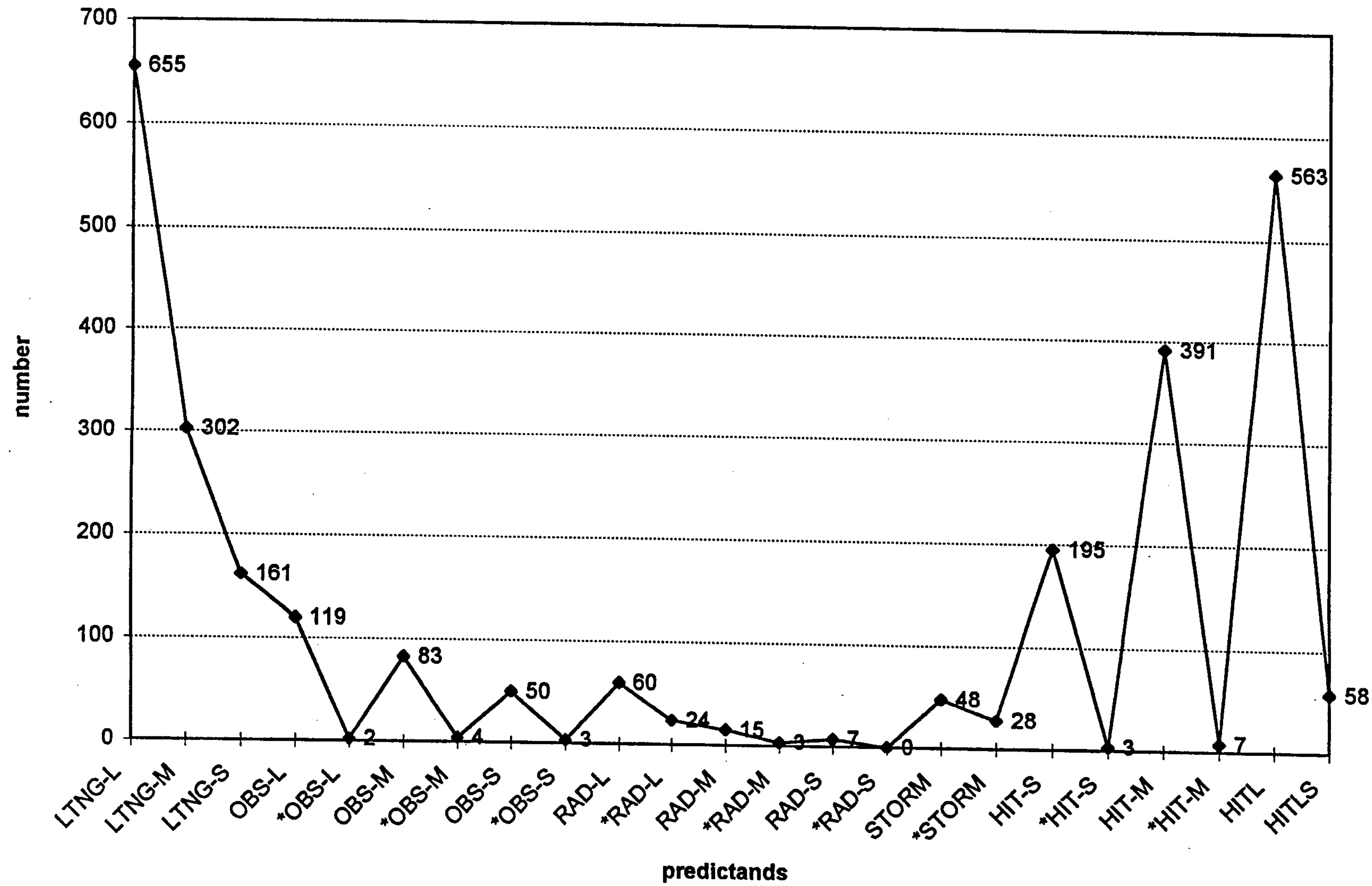


Figure 7. Same as in Figure 4 but for Region 3 (Southwest).

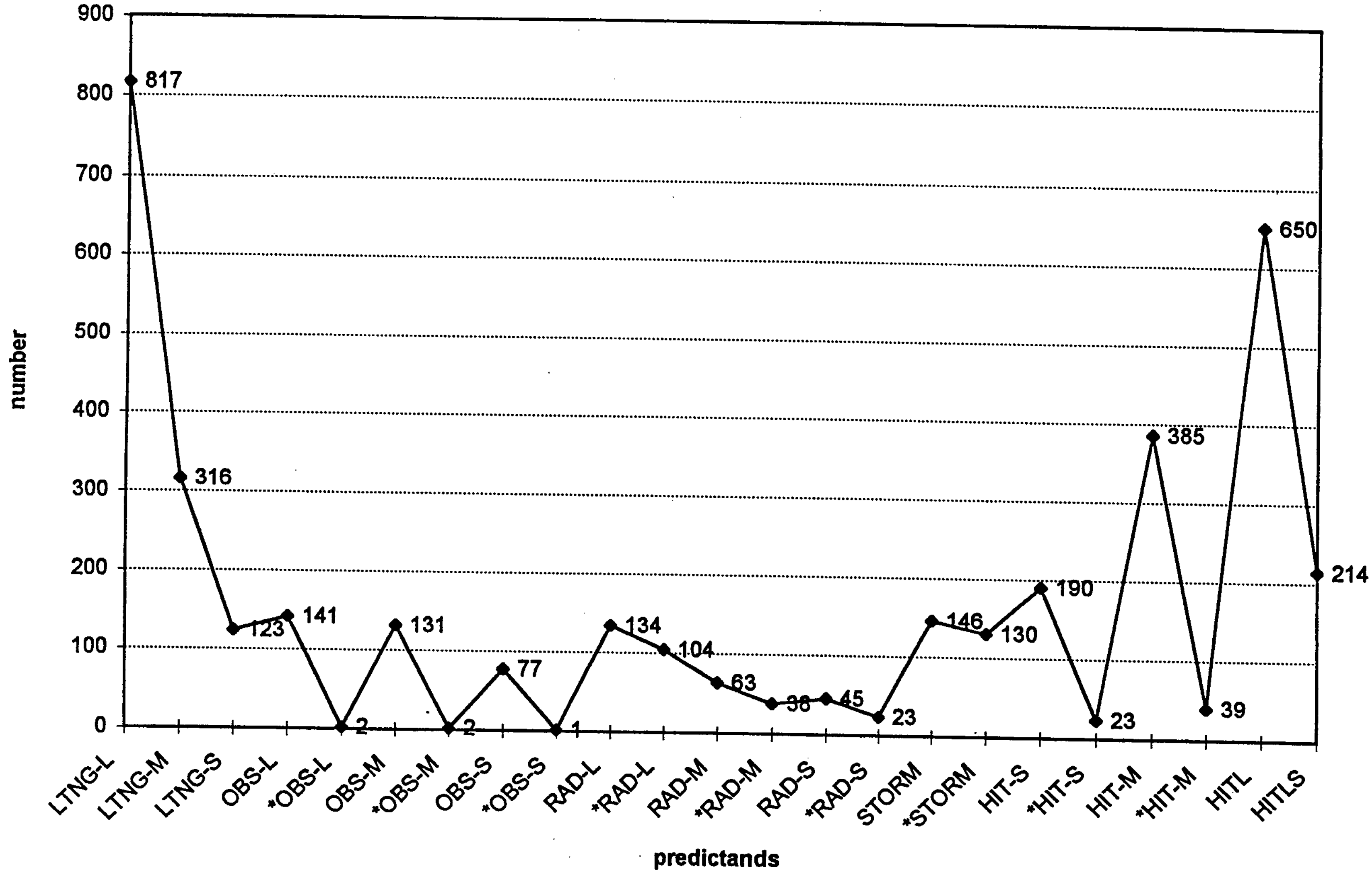


Figure 8. Same as in Figure 4 but for Region 4 (Panhandle)

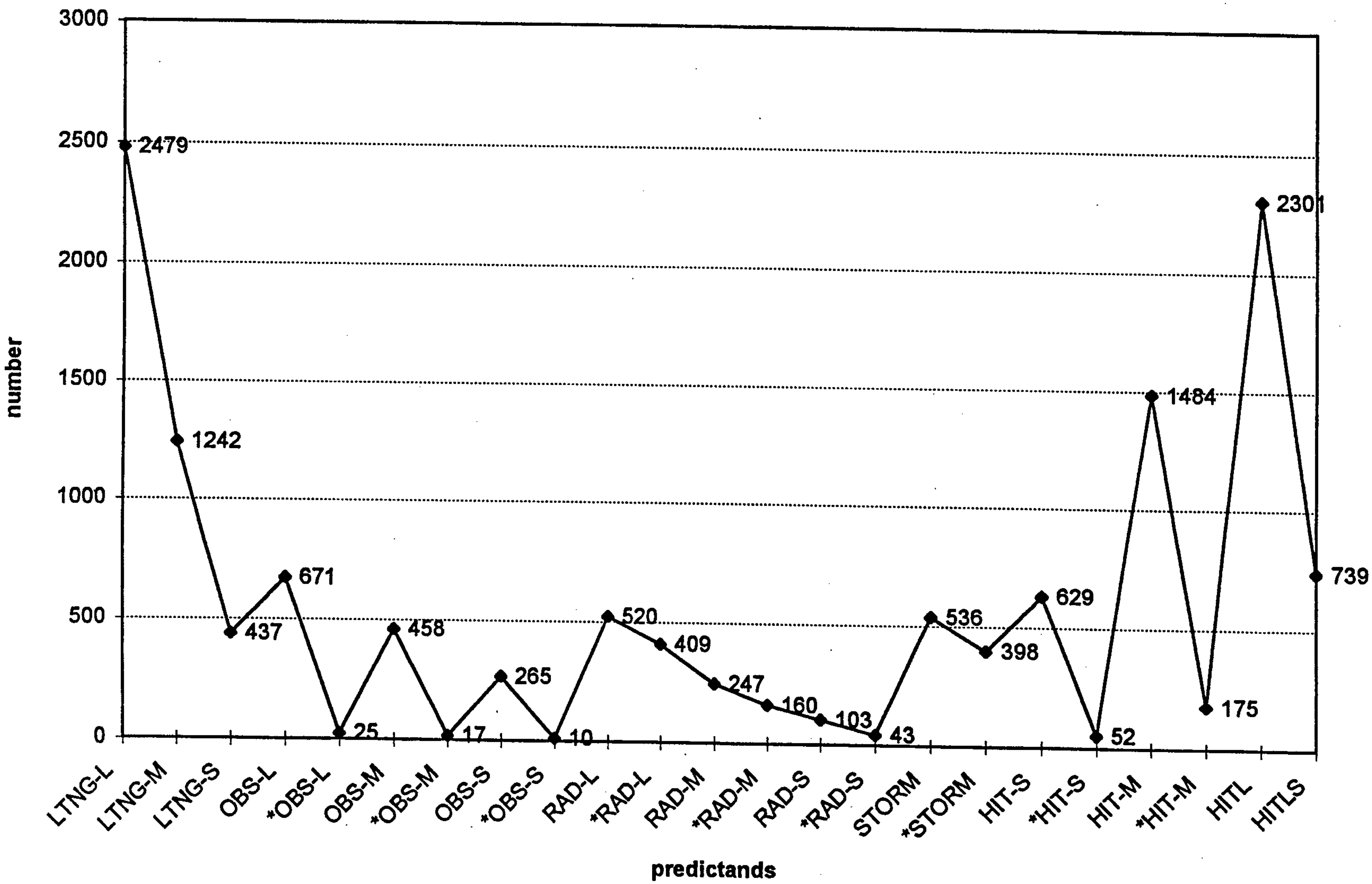


Figure 9. Same as in Figure 4 but for Region 5 (Midwest's Northern Tier)

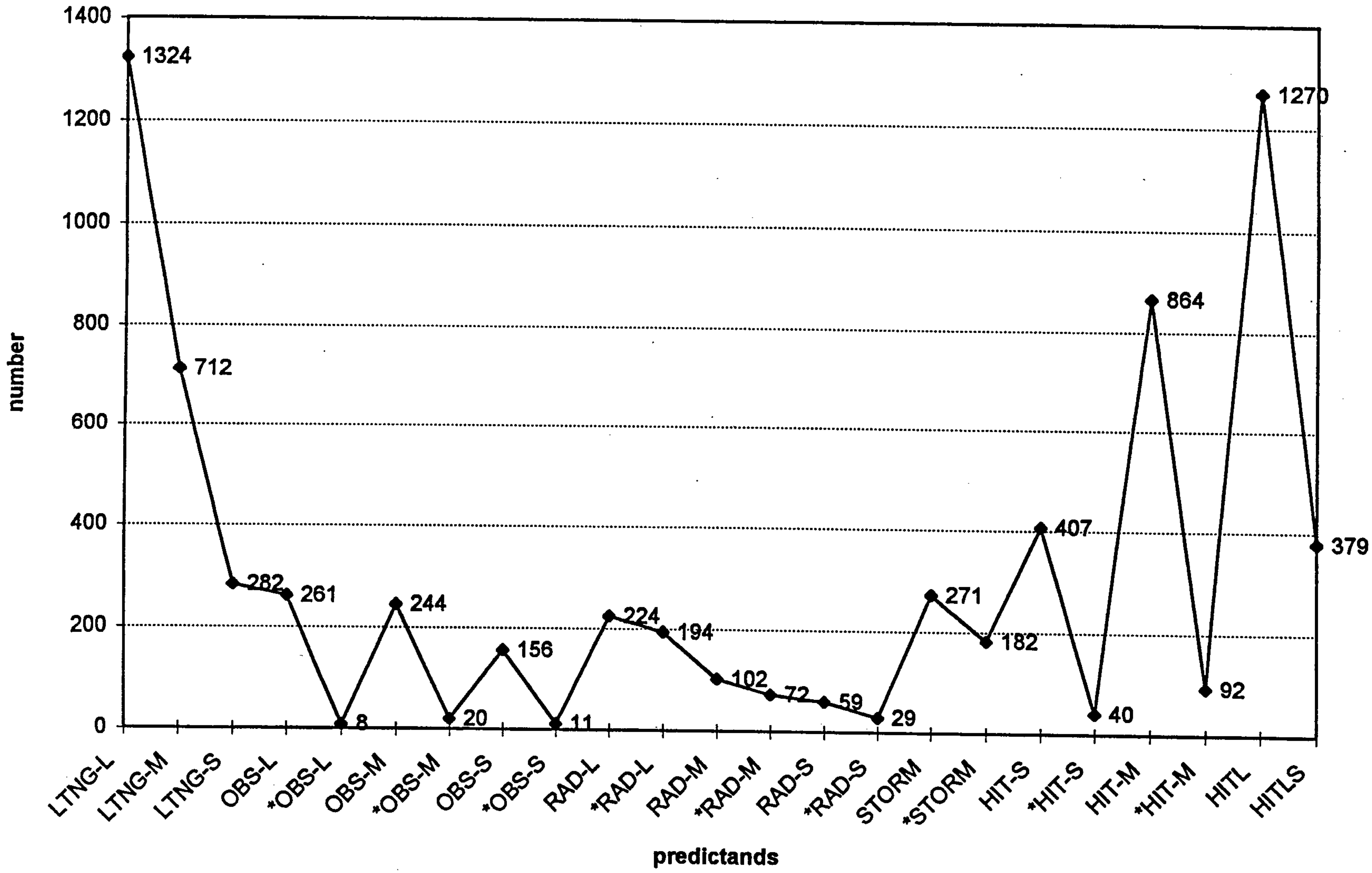


Figure 10. Same as in Figure 4 but for Region 6 (Southwest Interior).

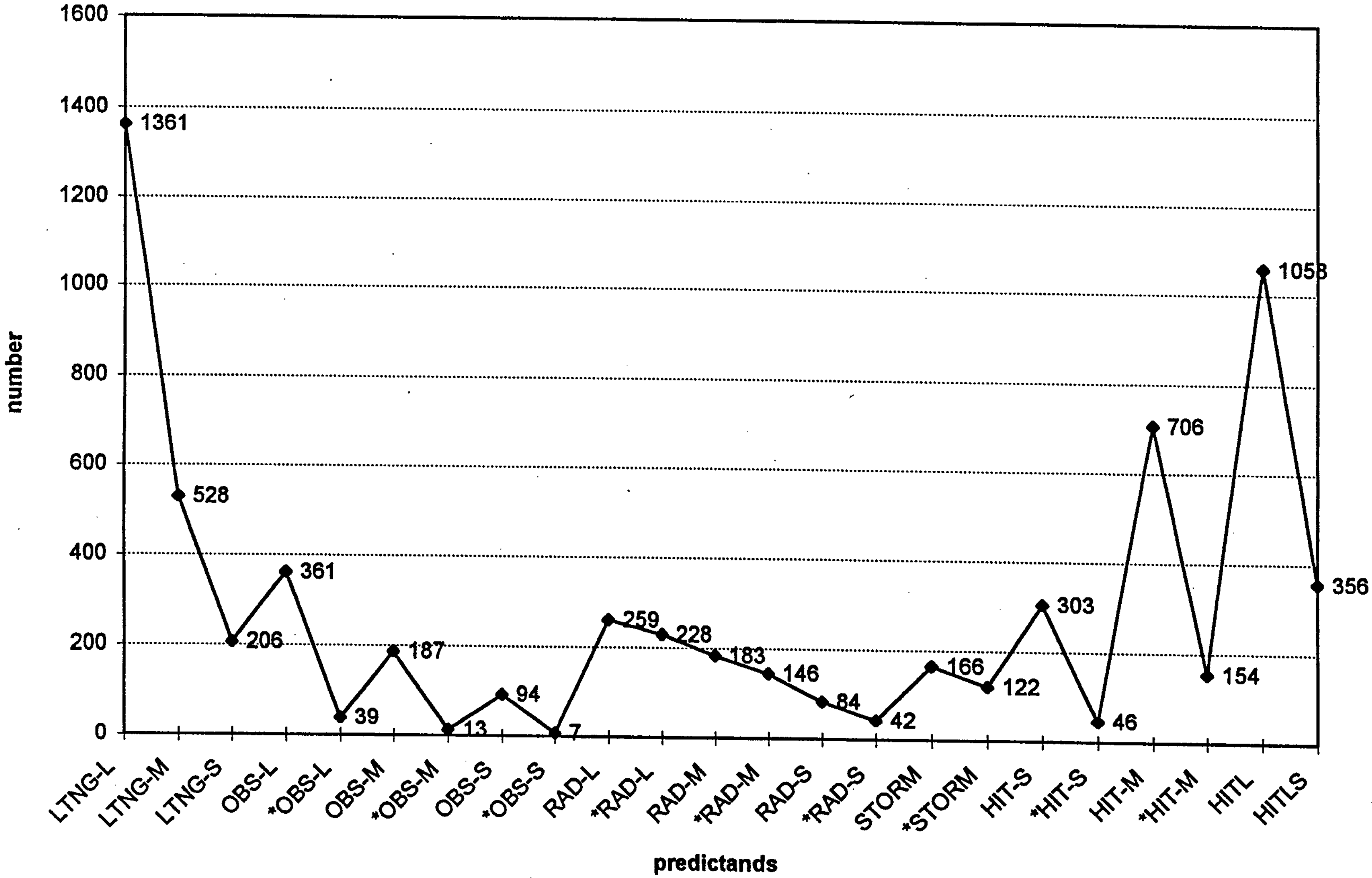


Figure 11. Same as in Figure 4 but for Region 7 (Gulf Coast).

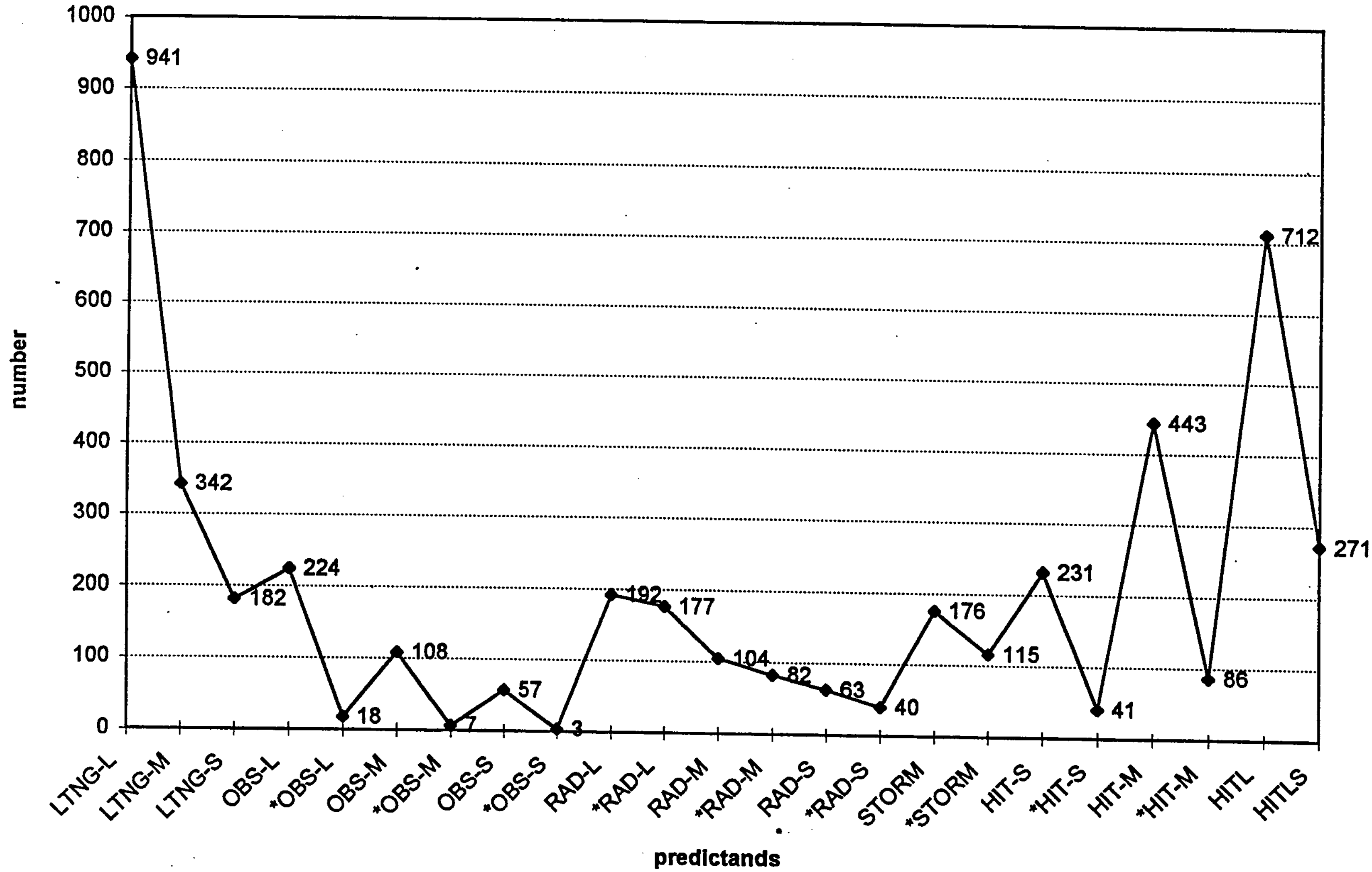


Figure 12. Same as in Figure 4 but for Region 8 (Southeast Coast).

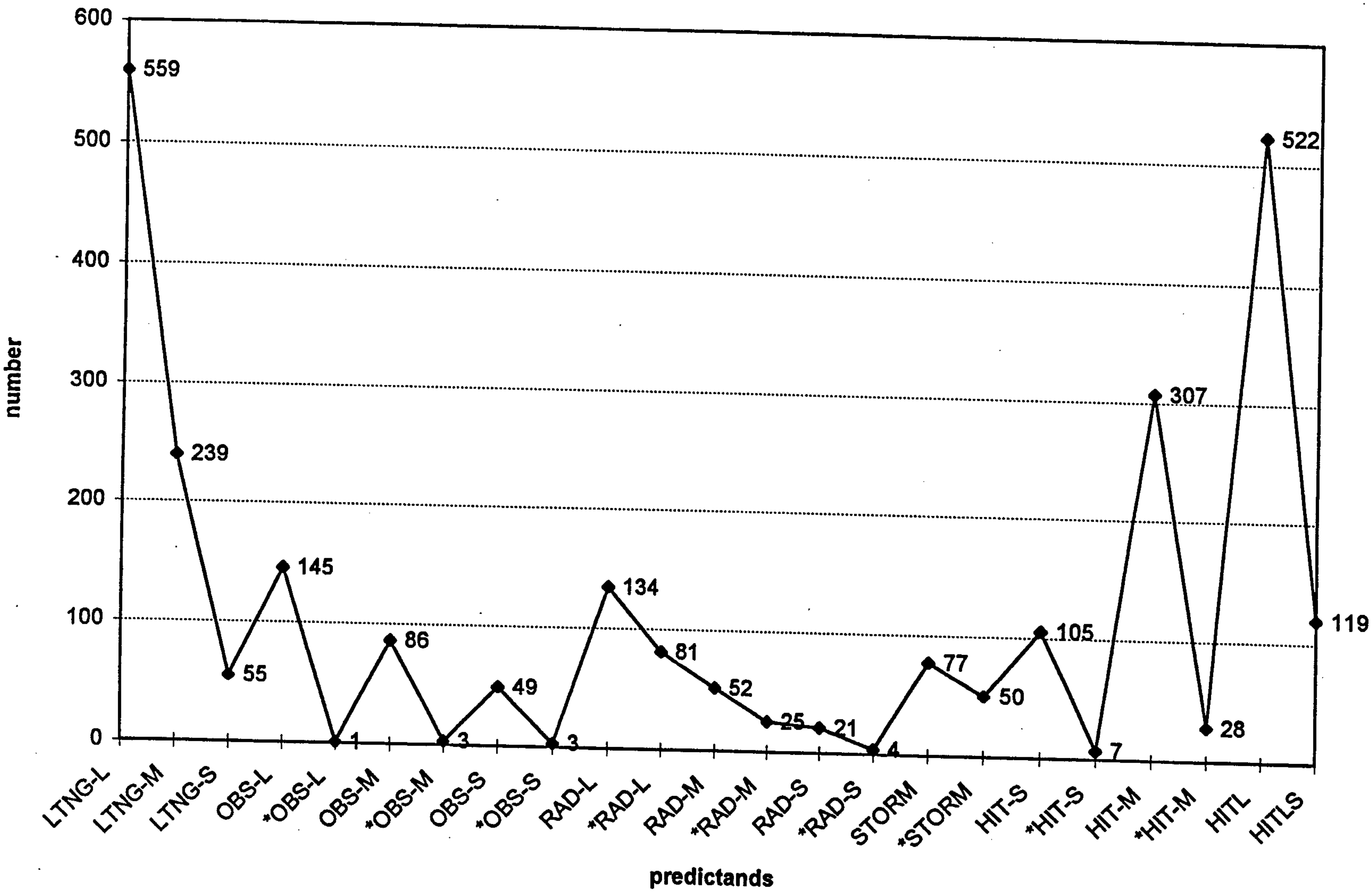


Figure 13. Same as in Figure 4 but for Region 9 (Great Lakes)

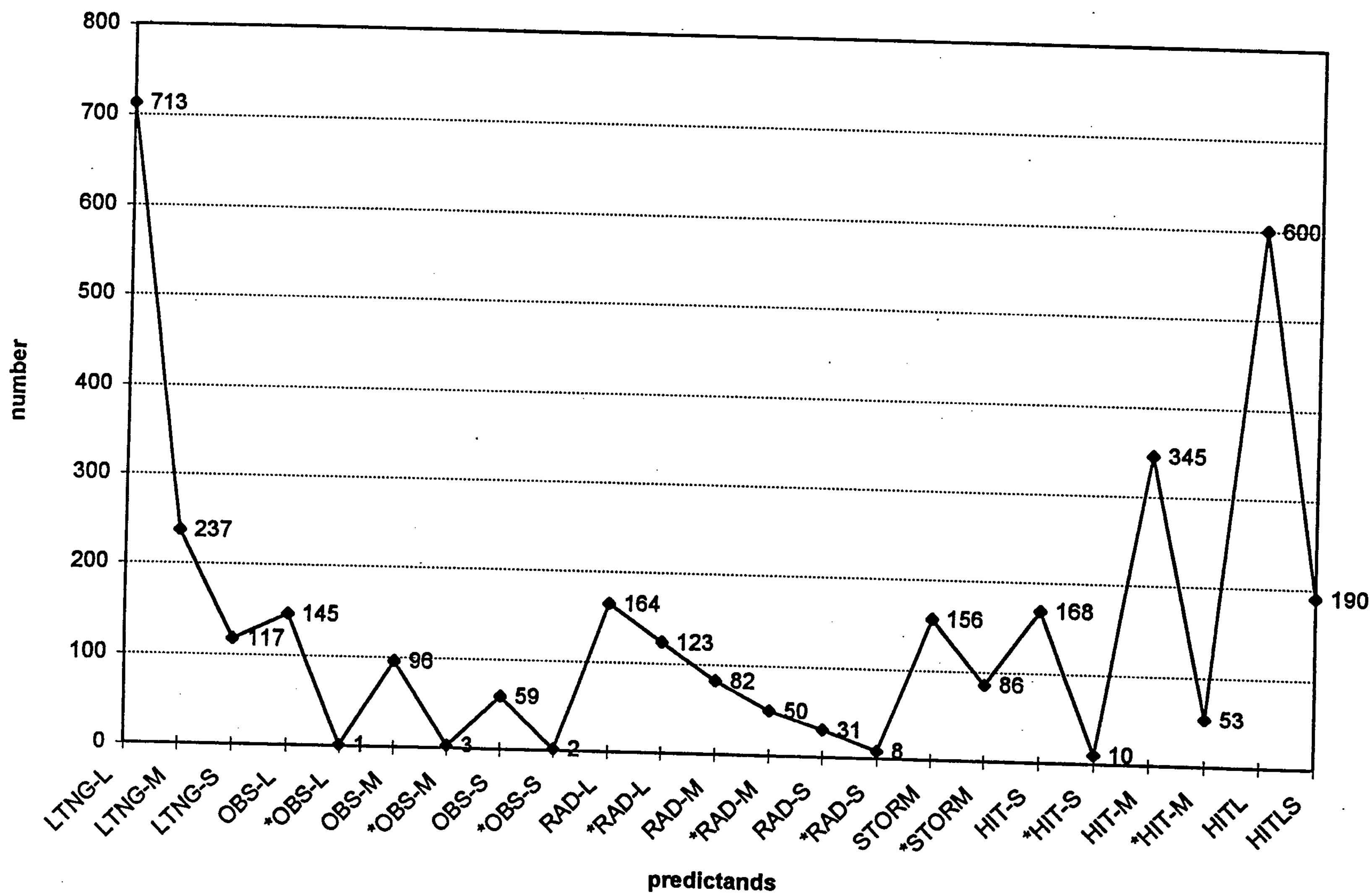


Figure 14. Same as in Figure 4 but for Region 10 (Appalachians).

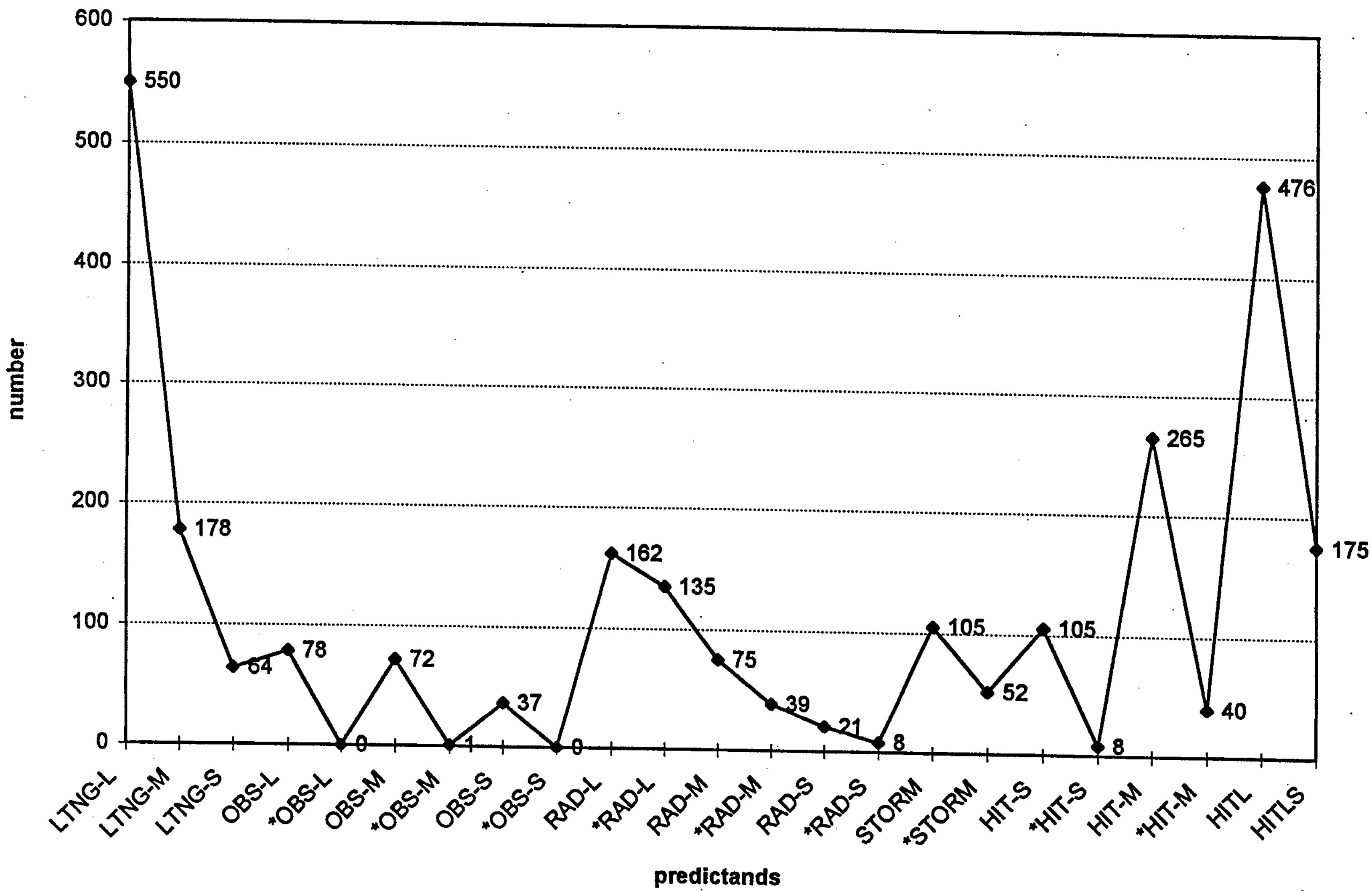


Figure 15. Same as in Figure 4 but for Region 11 (Northeast Coast).

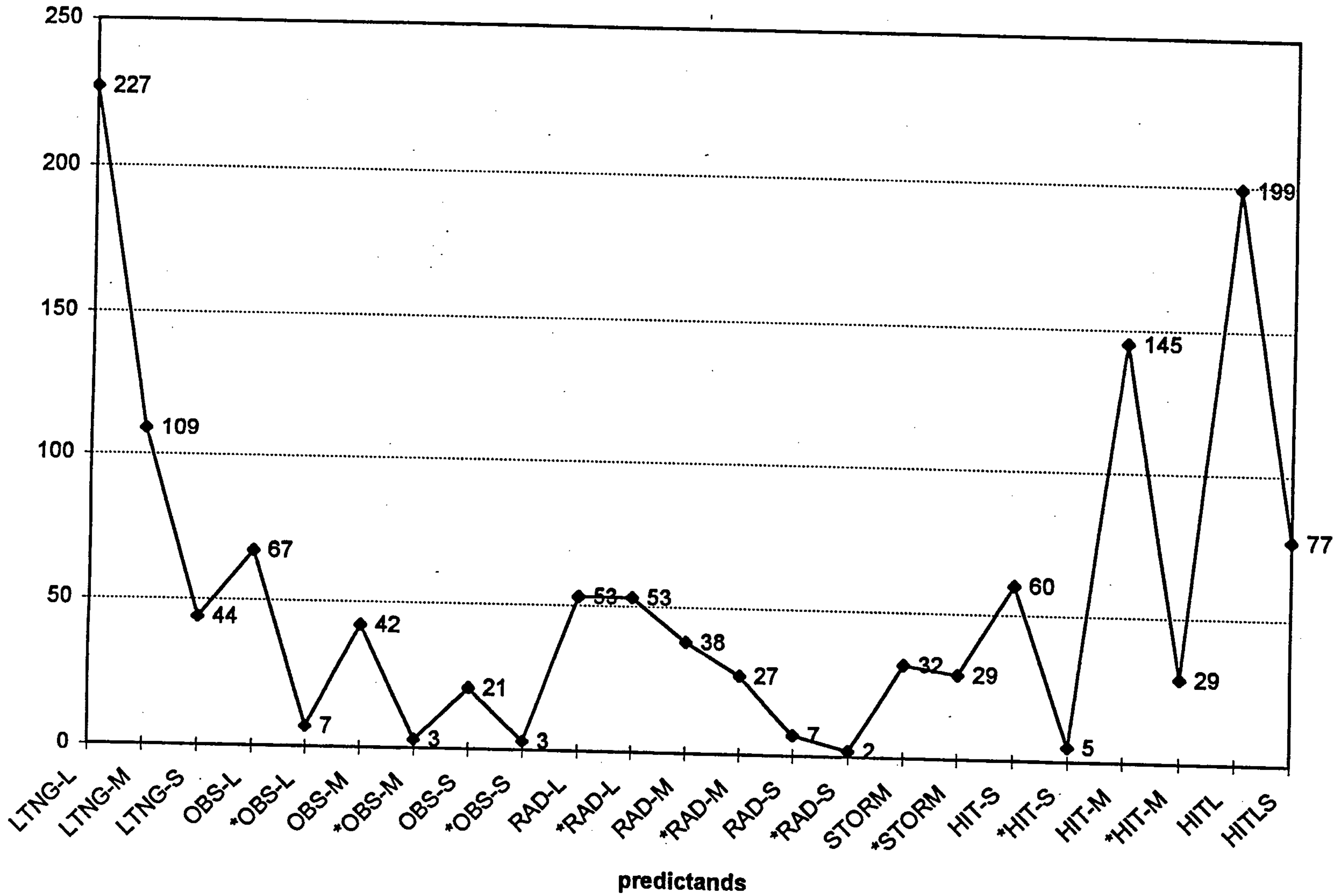


Figure 16. Same as in Figure 4 but for Region 12 (Southeast).

3.1 Thunderstorm Occurrence

All thunderstorm and severe storm occurrence are the predictands, and the types of data used to construct the set of predictands are shown in Table 4. Severe storms were characterized as storms exhibiting funnel clouds or tornadoes, hail 0.75 in. or larger, wind 50 knots or greater, and/or heavy rain.

Table 4. Types of Data Used to Generate Predictands.

	Severe + Nonsevere	Severe
Surface Observations	yes	yes
Manually-digitized radar (MDR)	yes*	yes*
Lightning	yes	no
STORM SUMMARY	yes	yes

*Only available for 1993.

3.1.1 Manually Digitized Radar (MDR) Data

Much less MDR data was available for this study than expected. MDR data was available for approximately 20 percent of the days in the period of interest. However, the agreement with lightning data was excellent, and so we decided to use as much of the data as was available to us. Severe storms were defined as those with MDR pixels exhibiting rainfall rate echo intensities of strong, very strong, intense, or extreme (digital video integrator and processor (DVIP) level of 3 or greater). Storm activity in general was defined as DVIP levels 1-6 inclusive (we made a judgment call that DVIP level 1, weak intensity, indicated thunderstorm occurrence).

3.1.2 Surface Observations

Station observations were obtained in DATSAV2 surface data fixed length format and analyzed for any type of storm activity. Severe storms were those with weather type encoded as 19 (funnel/tornado), 27 (hail), 97 (heavy tstm w/o hail), 99 (heavy tstm w/hail), and/or wind speed of 263 tenths of meters per second or greater (indicating wind speed of 50 knots or greater).

3.1.3 Lightning Flash Data

We relied most heavily on Global Atmospheric lightning flash data, taken over the National Lightning Detection Network and obtained through AFCCC, OL-A, for detection of storm activity. It was formatted on OL-A data field structure, and contained an overwhelming amount of detailed flash data, which had to be edited for our use. An efficient software parsing and editing routine was employed to reduce this voluminous data set. The area encompassed by the lightning predictand was in the shape of a square not a circle, and this represents what we estimate as the greatest source of error in this study.

3.1.4 National Climatic Data Center (NCDC) Storm Data

NCDC storm data was reviewed manually by county and state and represents the most reliable source of data for severe storm detection. Since we were initially relying on both MDR data and Storm Data for severe storm identification, our discovery of the MDR data shortage left us with about half of the severe storm information we expected to have.

We pooled predictands into 3 types, to help set the stage for eventual evaluation of predictions using model forecast data. Ostensibly, the area of storm detection should equal model grid size; however, we also produced two other predictand sets for comparison purposes.

The predictand types are as shown in Table 5, and represent both a circular areal extent around RAOB sites and a temporal period about RAOB synoptic times (00Z and 12Z). The spatial scheme of 100 statute miles (s.m.) is similar in area to that used by Charba (1979)[6]. A radius of 100 s.m. is roughly equivalent to one-half the average spacing of upper air observation sites across the CONUS (one- half of 400 km) or 144 s.m. (Maddox, 1980)[7]. Figure 17 shows "L" areal coverage around each RAOB site. Figure 18 shows the temporal coverage about RAOB launch time. Last, Table 6 shows the number of storm-related RAOBs, both severe and overall (including severe).

Table 5. Predictand Types

	S(mall)	M(edium)	L(arge)
Predictand "Type"			
Detection w/i circle of radius & (statute miles):	33	66	100
Detection w/i ± h (hours):	1	3	6

Table 6. Number (Percent) of Storm-Related RAOBs

Region	1992-1993 Overall/Total	%	% of Total	1992-1993 Severe/Total	%	% of Total	1992-1993 Severe/Storms	%
1	93/1355	(6.8)	0.4	8/1355	(.6)	0.03	8/93	(8.6)
2	1166/3595	(32.4)	4.9	89/3595	(2.5)	0.38	89/1166	(7.6)
3	563/1447	(38.9)	2.4	58/1447	(4.0)	0.24	58/563	(10.3)
4	650/1346	(48.3)	2.8	214/1346	(15.9)	0.91	214/650	(32.9)
5	2301/4836	(47.6)	9.8	739/4836	(15.3)	3.10	739/2301	(32.1)
6	1270/2373	(53.5)	5.4	379/2373	(15.9)	1.50	379/1270	(29.8)
7	1058/2048	(51.7)	4.5	356/2048	(17.4)	1.50	356/1058	(33.6)
8	712/1384	(51.4)	3.0	271/1384	(19.5)	1.15	271/712	(38.1)
9	522/1697	(30.7	2.2	119/1697	(7.0)	0.51	119/522	(22.7)
10	600/1674	(35.8)	2.6	190/1674	(11.4)	0.81	190/600	(31.6)
11	476/1339	(35.5)	2.0	175/1339	(13.1)	0.74	175/476	(36.7)
12	- 199/336	(59.2)	0.8	77/336	(22.9)	0.33	77/199	(38.6)
Total	9610/23430	(41.0)		2675/23430	(11.4)		2675/9610	(27.8)

3.2 Maximum Storm Cloud Top Height

Equilibrium Level (EL) height was used as a predictand to estimate maximum storm cloud top height. EL is the height where the temperature of a buoyantly rising parcel again becomes equal to the temperature of the environment. This predictand was developed using only the "L" predictand data pool, including both non-severe and severe storms.

No additional data processing was necessary for this part of the study, because EL height had been one of the predictors for storm occurrence. Using EL height as a predictand meant using a non-binary predictand and ignoring those soundings not characterized by an EL (or where EL height is zero).

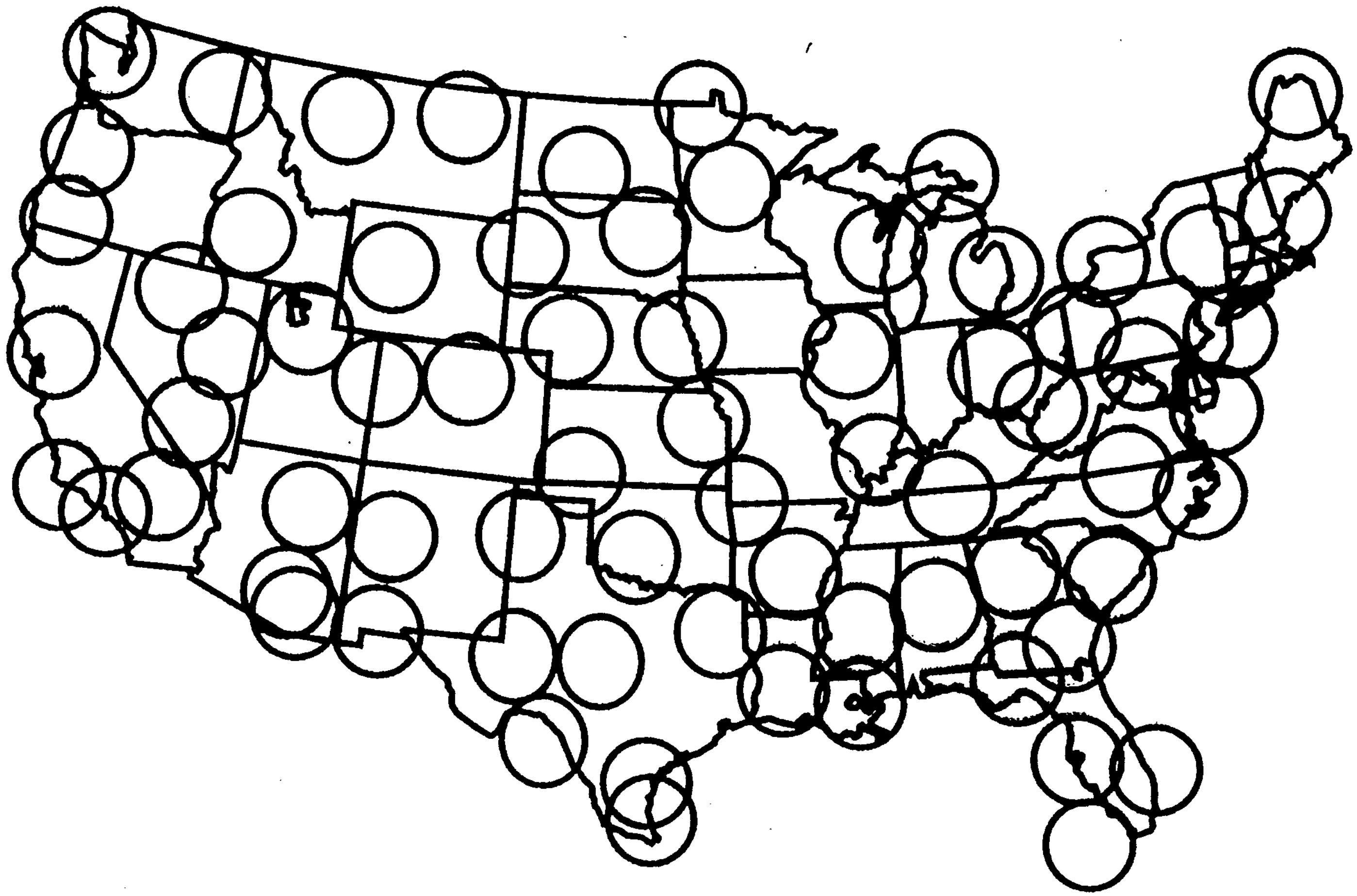


Figure 17. Spatial Coverage Around Each RAOB Station Site for the "L" Predictors.

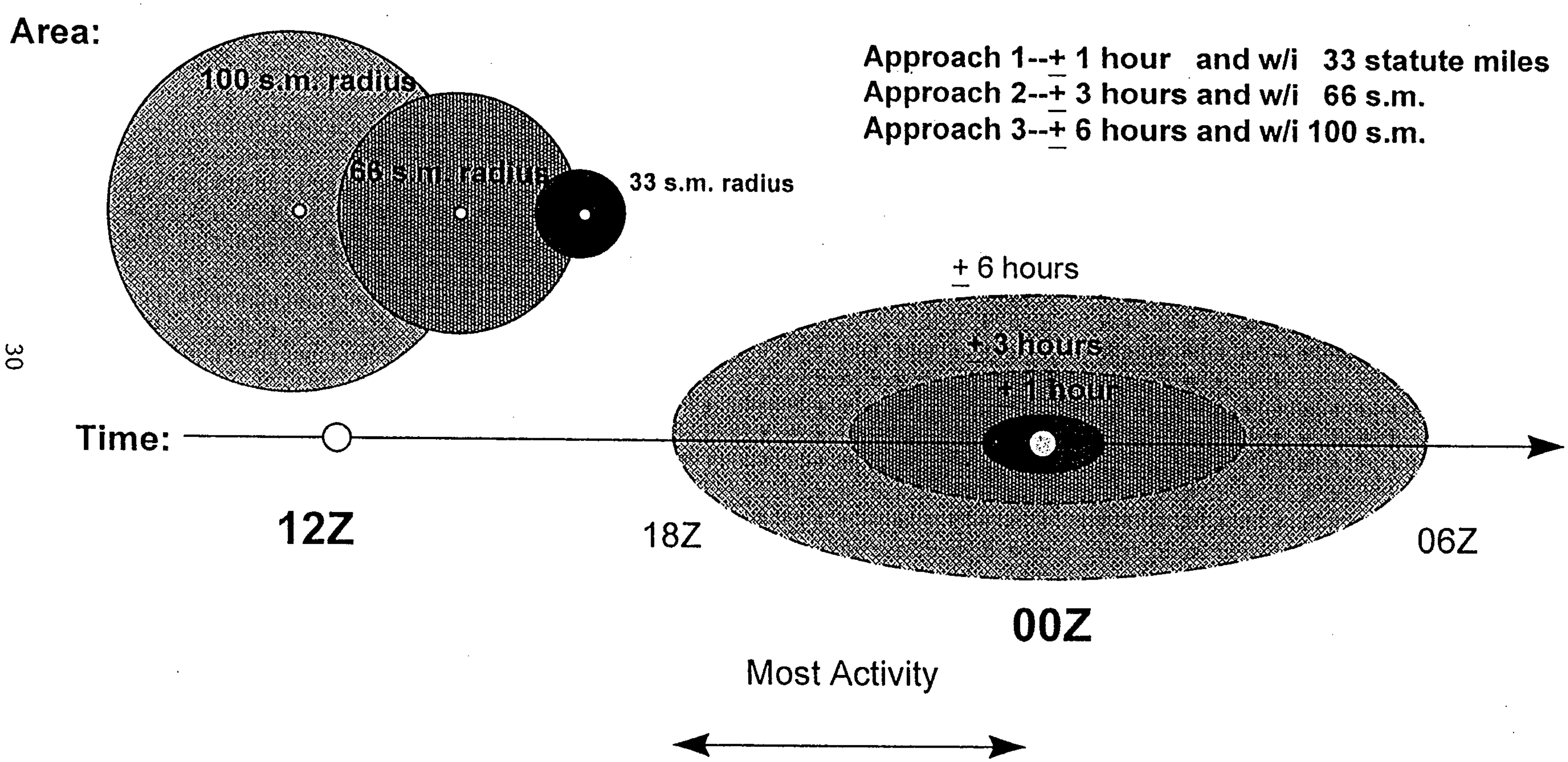


Figure 18. Temporal Coverage Used for Each RAOB Station Site.

4. CORRELATION COEFFICIENTS

Shown in Figures 19-57 are correlation coefficients for the top 36 predictors (all storm and severe storm cases) and each of the three predictand types, for each of the 12 CONUS regions and the CONUS overall. The equations presented in Section 6 were developed using the top ten predictors for the respective regions and predictand types and were identified using these charts. Although the predictors are listed in the same order on each chart, this is the SHARP-specified order in which predictors were databased. Note that one of the predictors, IN35, is not defined in this report–the authors of SHARP were not able to reconstruct its identity. It does appear to be useful in some cases, and if identified, we will make an addendum to this report.

This is a good place to note that when referring to these figures, one should be aware that not all totals are additive. For example, *STORM and *OBS-L are not mutually exclusive.

A striking characteristic are those charts showing predictors more highly correlated to severe storms than to severe and non-severe storms. This occurs markedly for Region 1-100 (Region 1, 100 s.m. predictand type) and Region 2-66, and more dramatically for Regions 4-33, 5-100, 6-100, 7-66, 8-66, 9-100, 10-100, 11-33, and 12-66. This may be partly explained by the relatively low number of RAOBs for several regions; the five regions with the lowest RAOB data pools (see Table 3) were Regions 1 (5.8%), 4 (5.8%), 8 (5.9%), 11 (5.7%), and 12 (1.4%). Figures 58 through 63 also shed some light. For Region 1, there were only 8 severe storms and 93 storms of any kind. The low number of Region 1 predictands is evident in Figures 58-63. Region 2 had 1166 out of 3595 storm-related RAOBs (32.4%), which was not particularly low but only 89 out of 3595 RAOBs that were severe storm-related (2.5%). Comparing Figures 58 and 59 illustrates this.

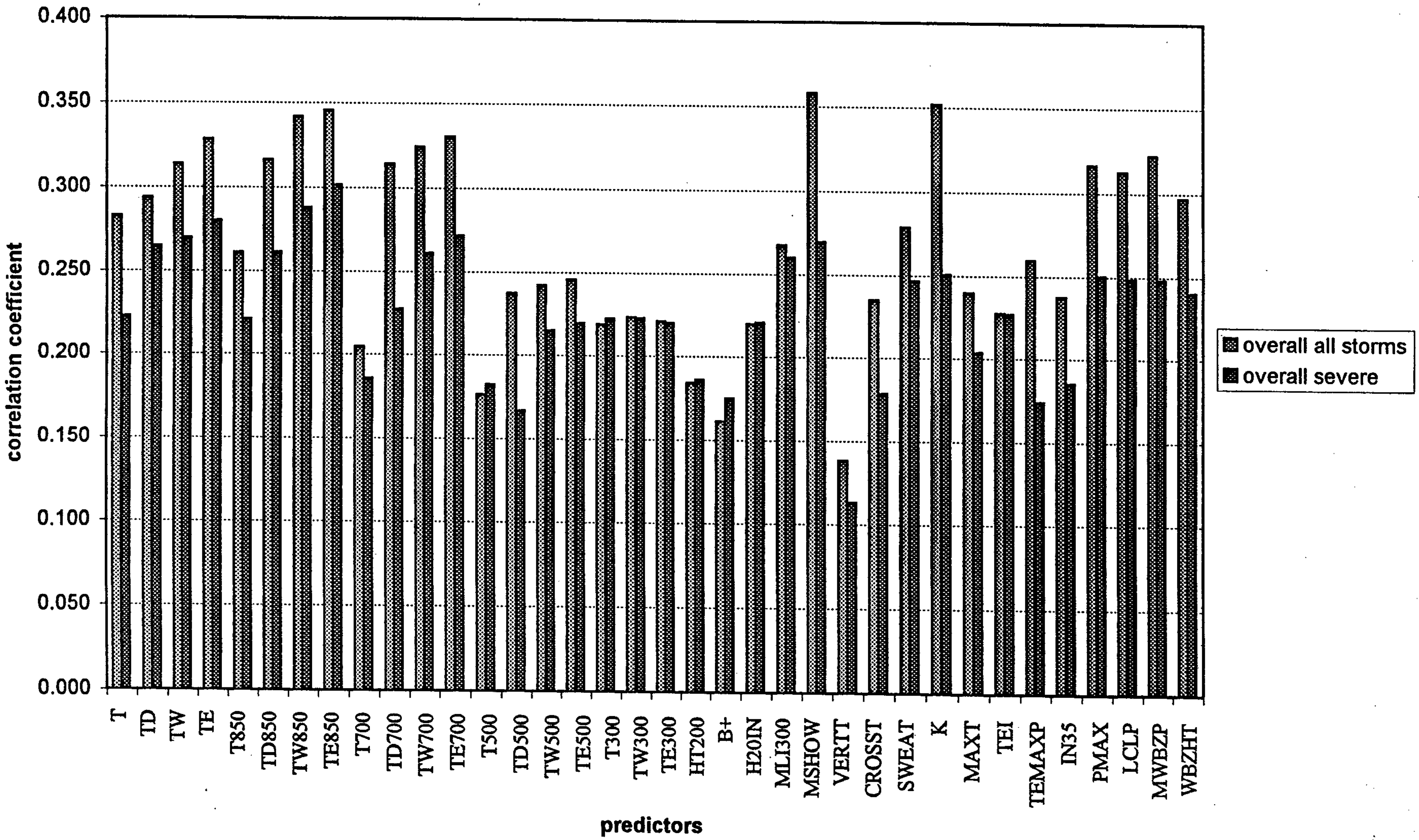


Figure 19. Correlation Coefficients for Overall Case Within 100 Statute Mile Radius.

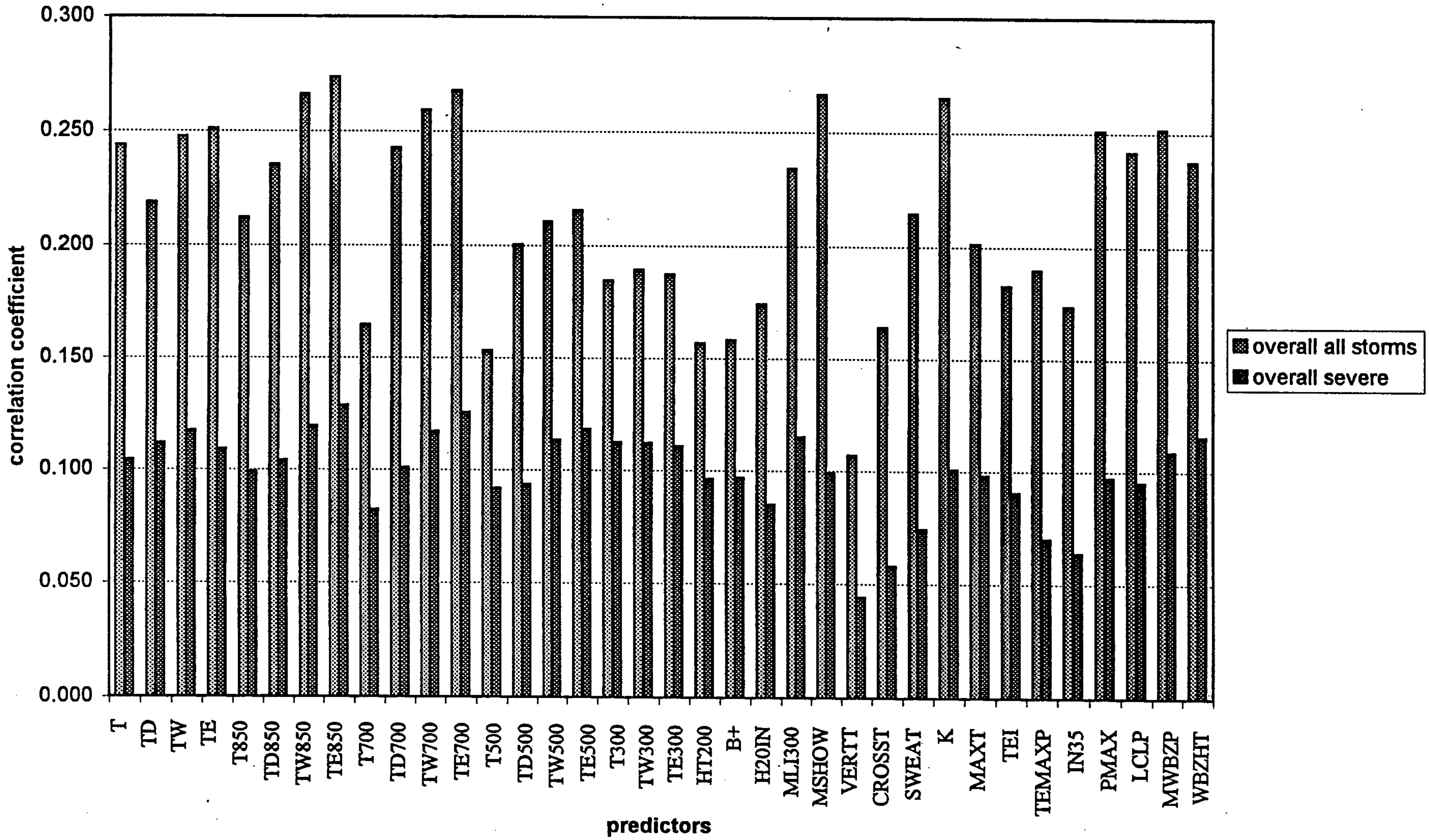


Figure 20. Same as Figure 19 but for 66 Statute Mile Radius.

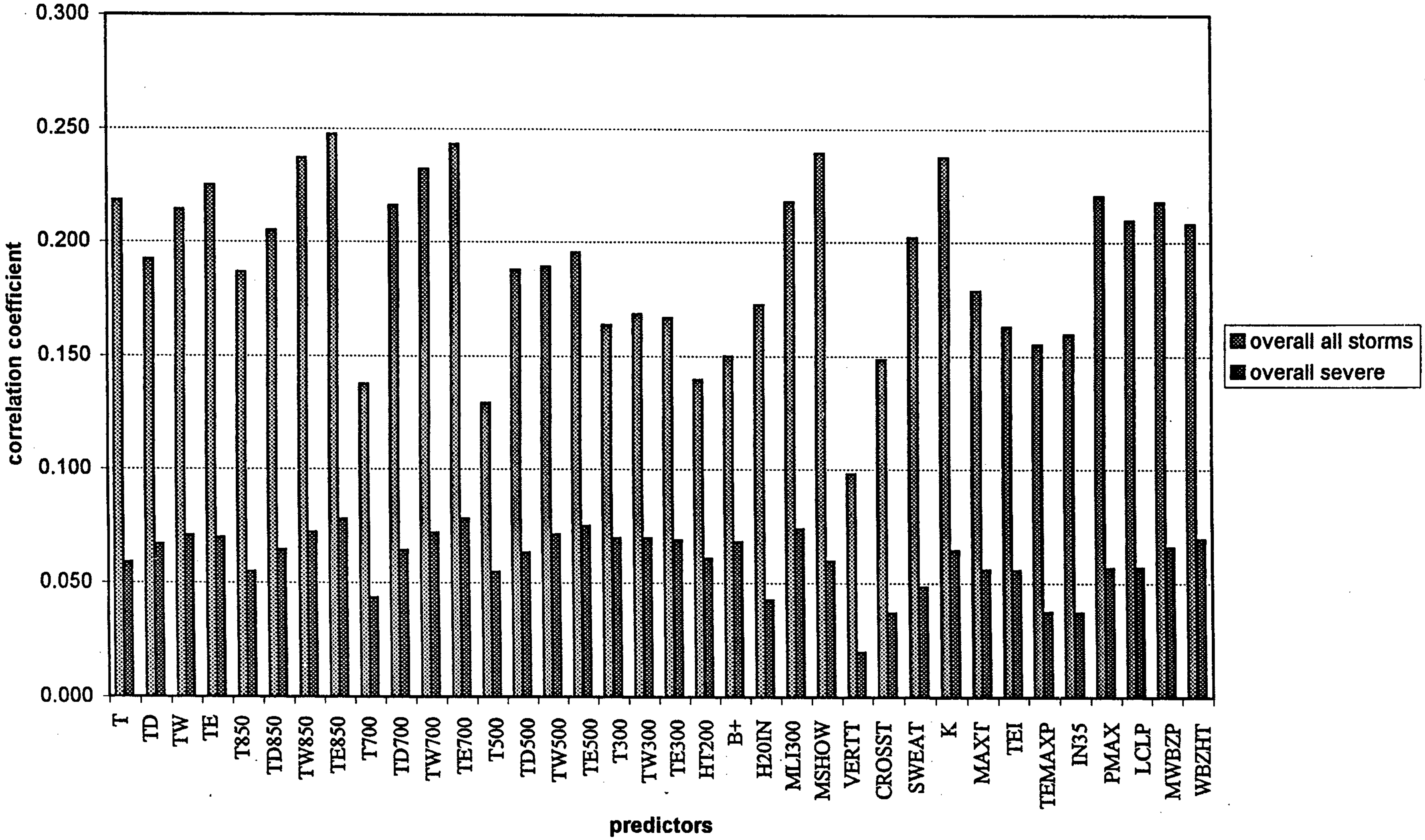


Figure 21. Same as Figure 19 but for 33 Statute Mile Radius.

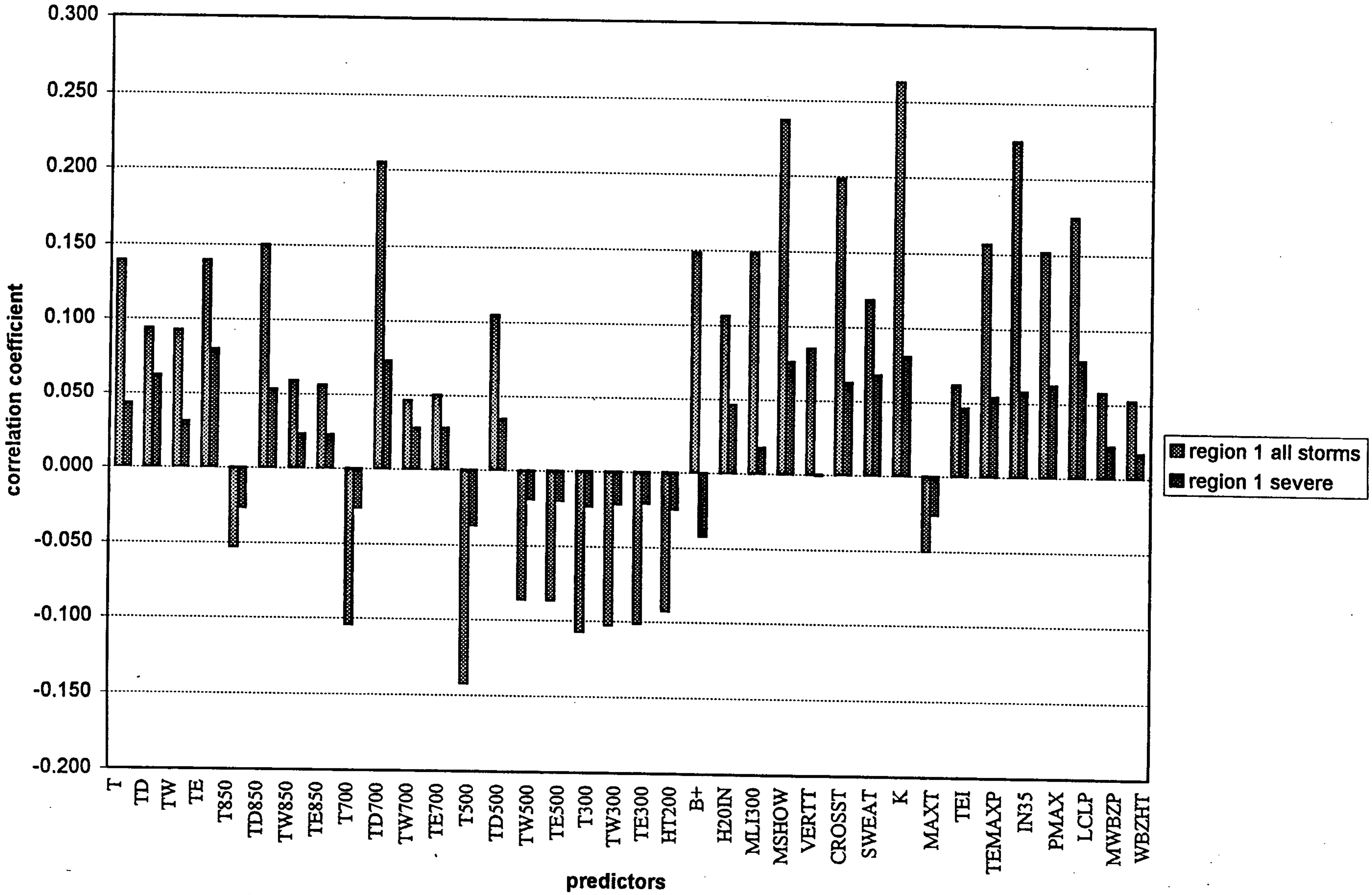


Figure 22. Correlation Coefficients for Region 1 (West Coast) Within 100 Statute Mile Radius.

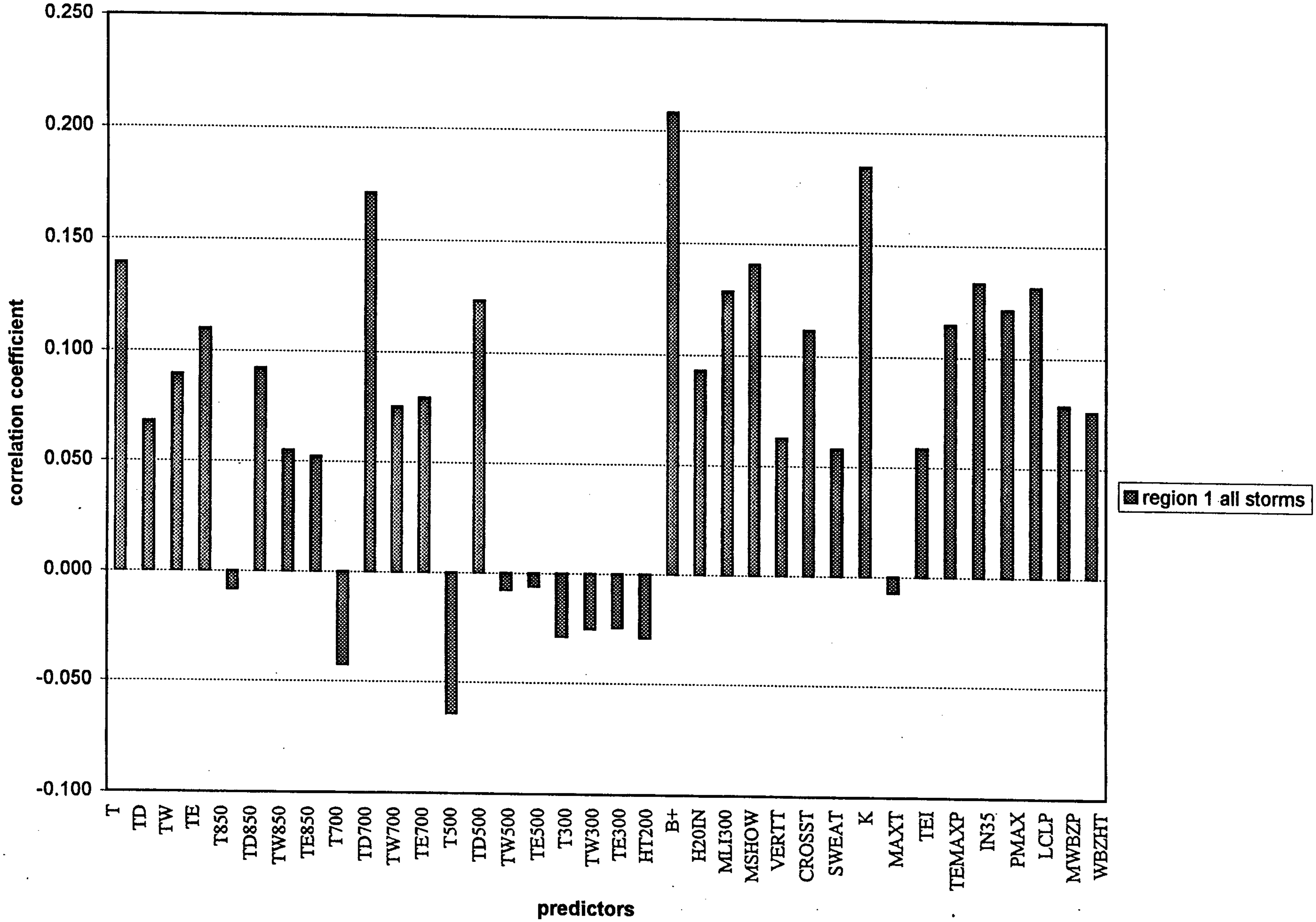


Figure 23. Same as Figure 22 but for 66 Statute Mile Radius.

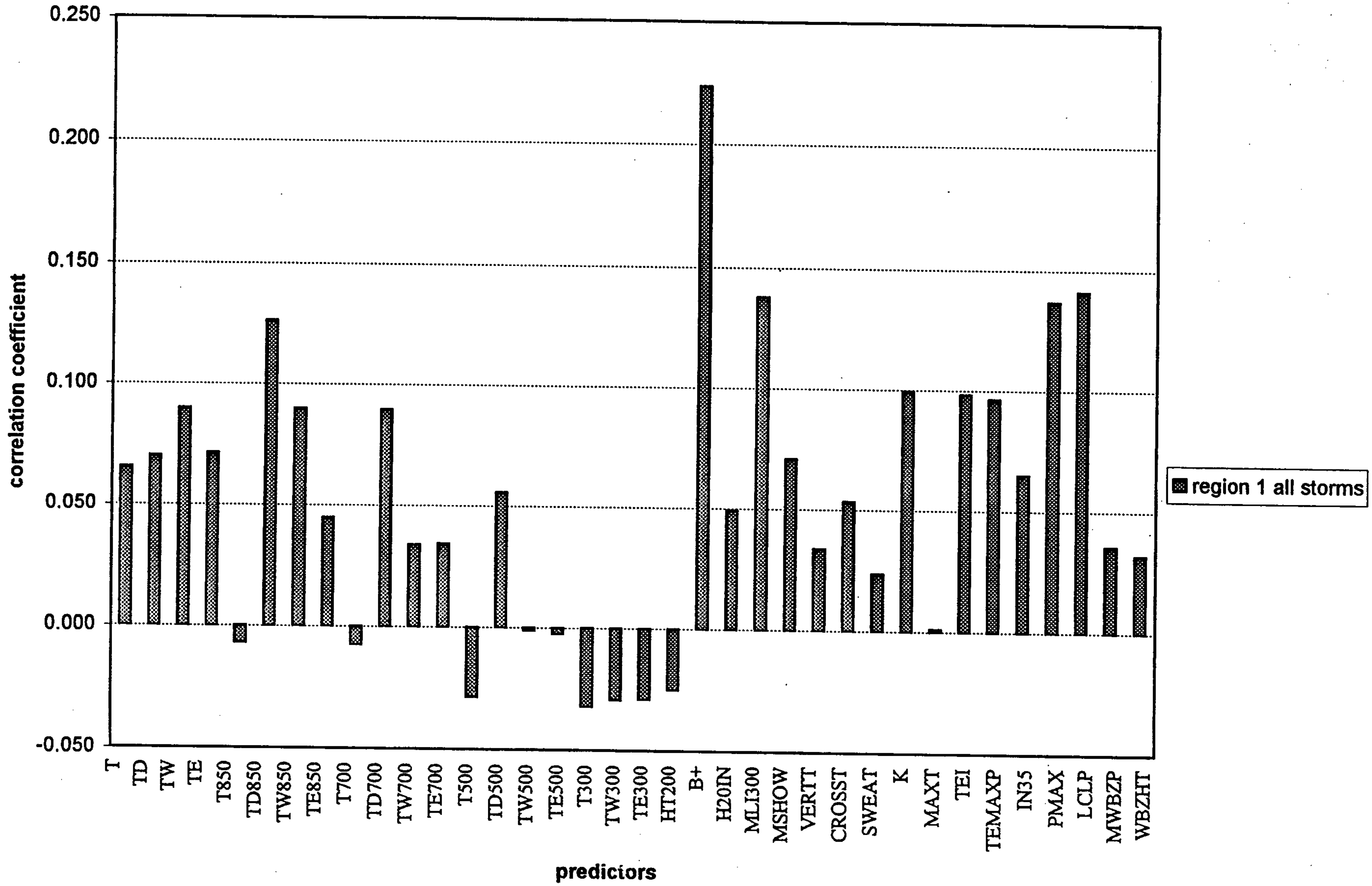


Figure 24. Same as Figure 22 but for 33 Statute Mile Radius.

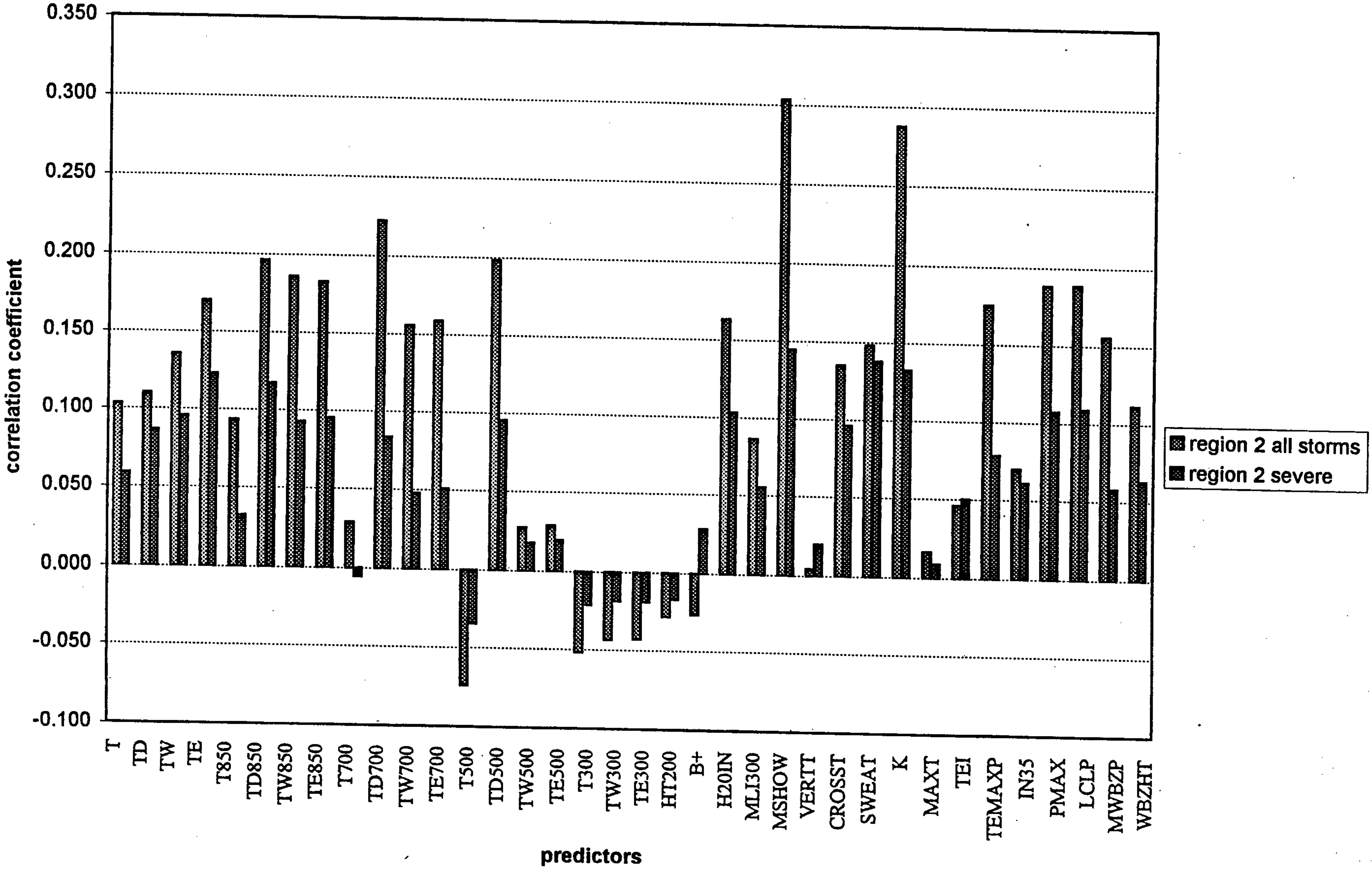


Figure 25. Correlation Coefficients for Region 2 (Intermountain West) Within 100 Statute Mile Radius.

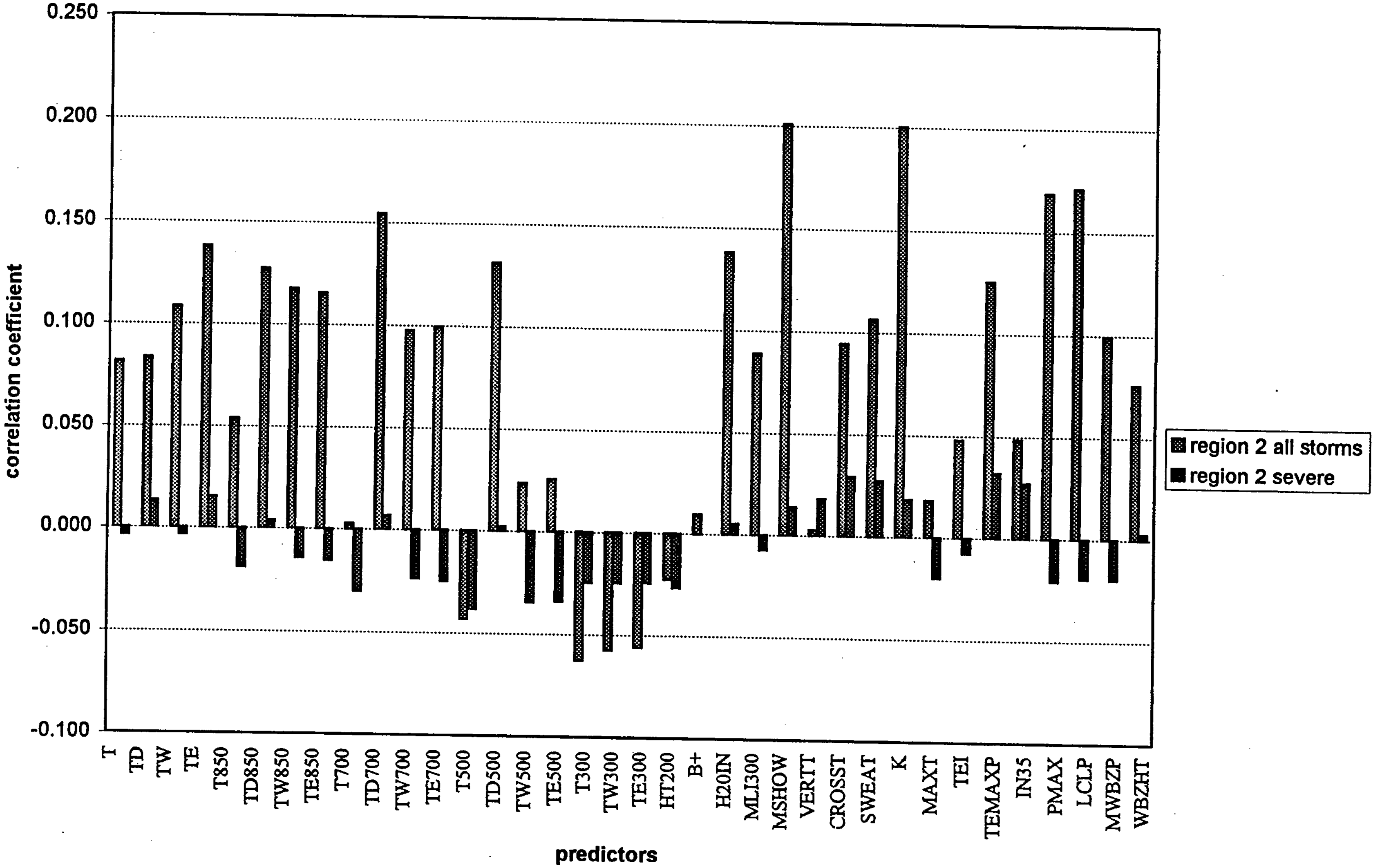


Figure 26. Same as Figure 25 but for 66 Statute Mile Radius.

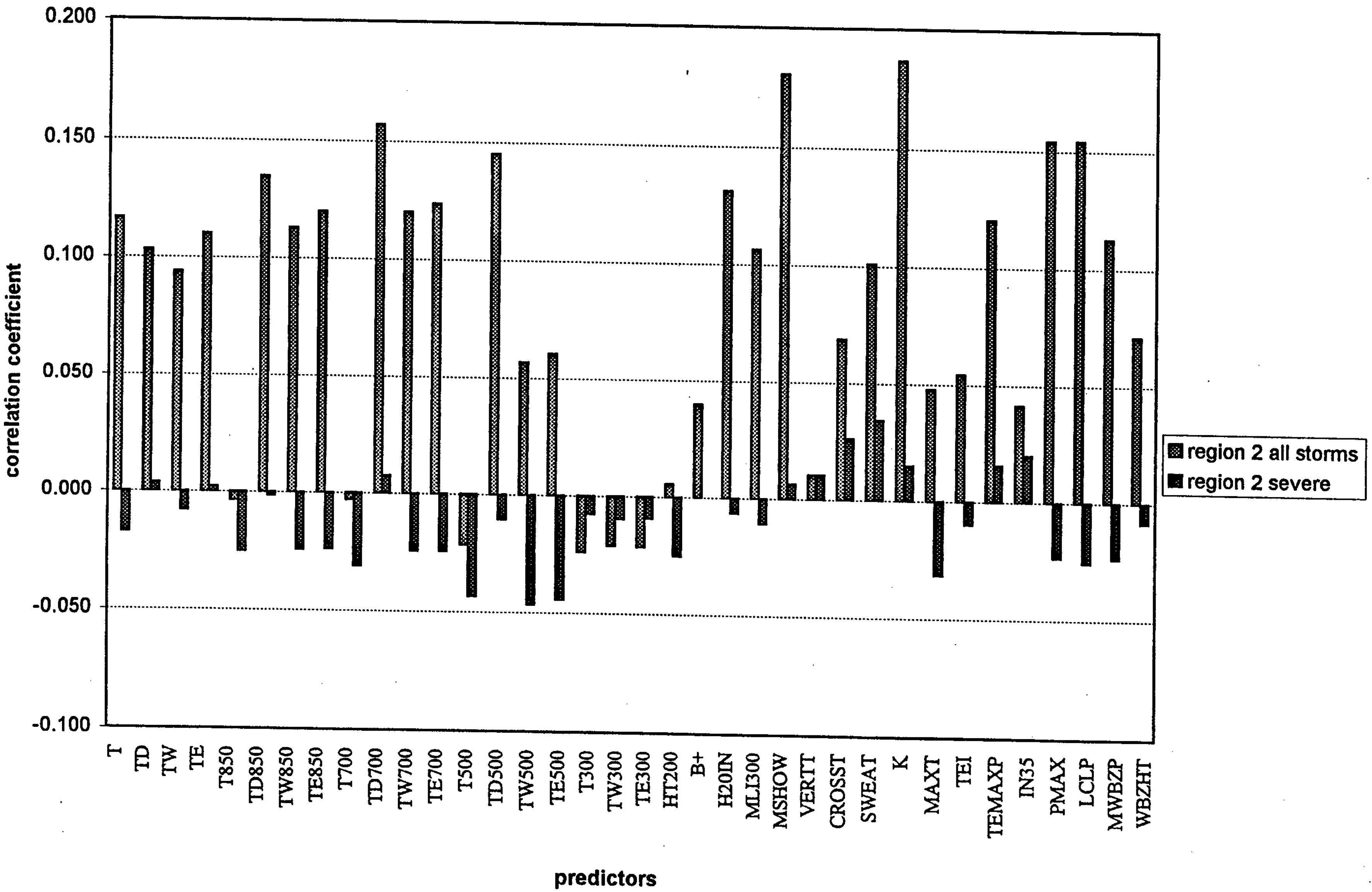


Figure 27. Same as Figure 25 but for 33 Statute Mile Radius.

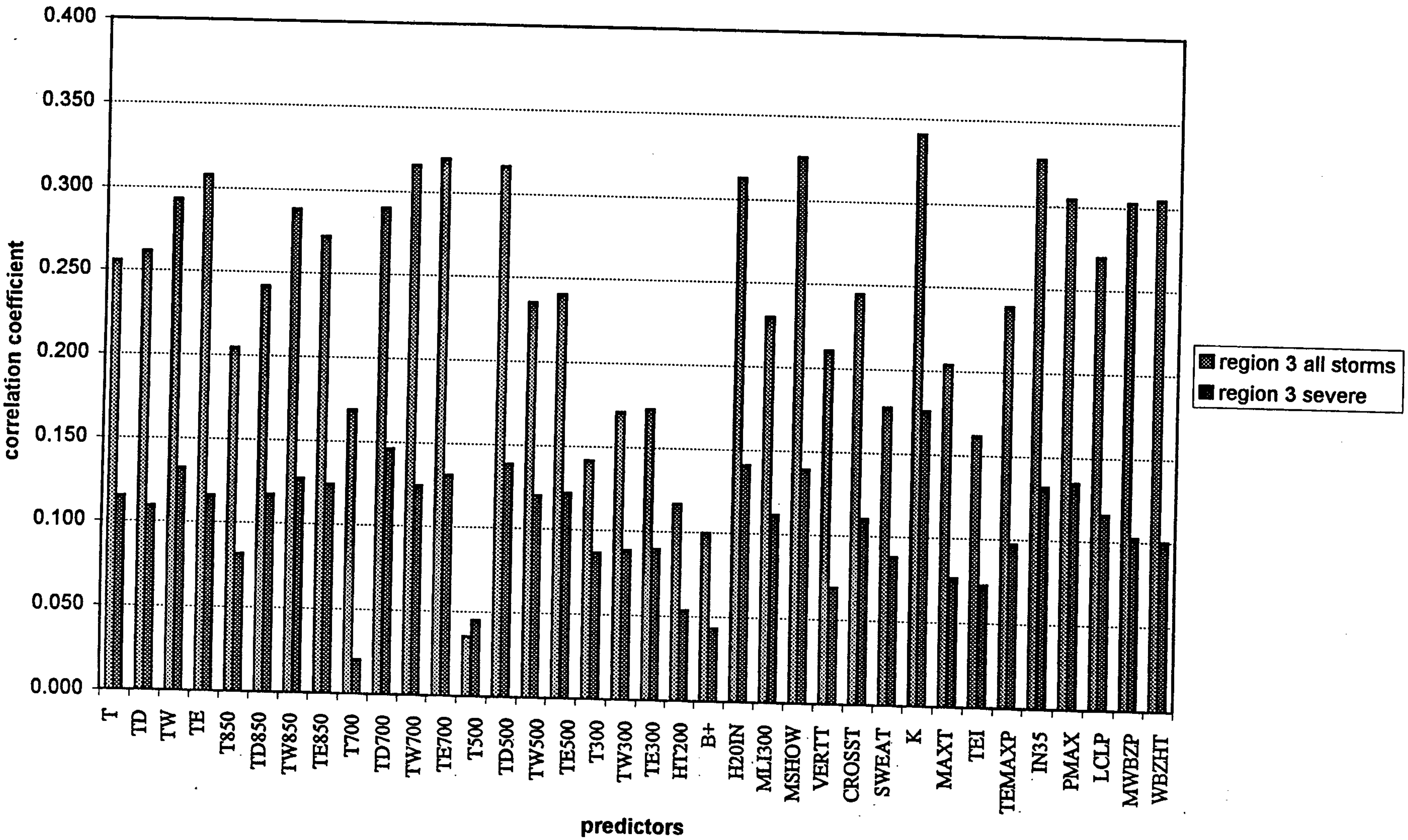


Figure 28. Correlation Coefficients for Region 3 (Southwest) Within 100 Statute Mile Radius.

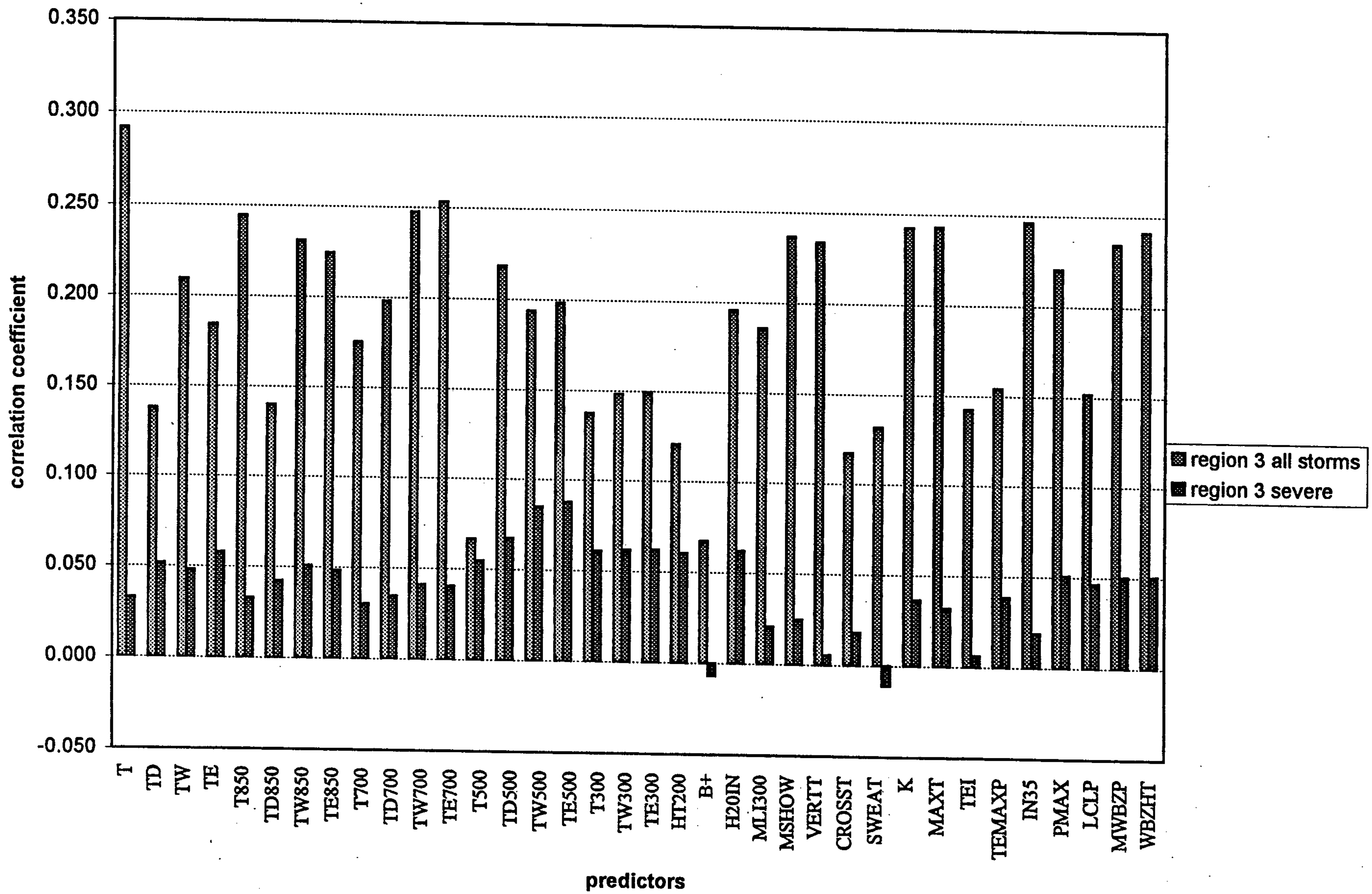


Figure 29. Same as Figure 28 but for 66 Statute Mile Radius.

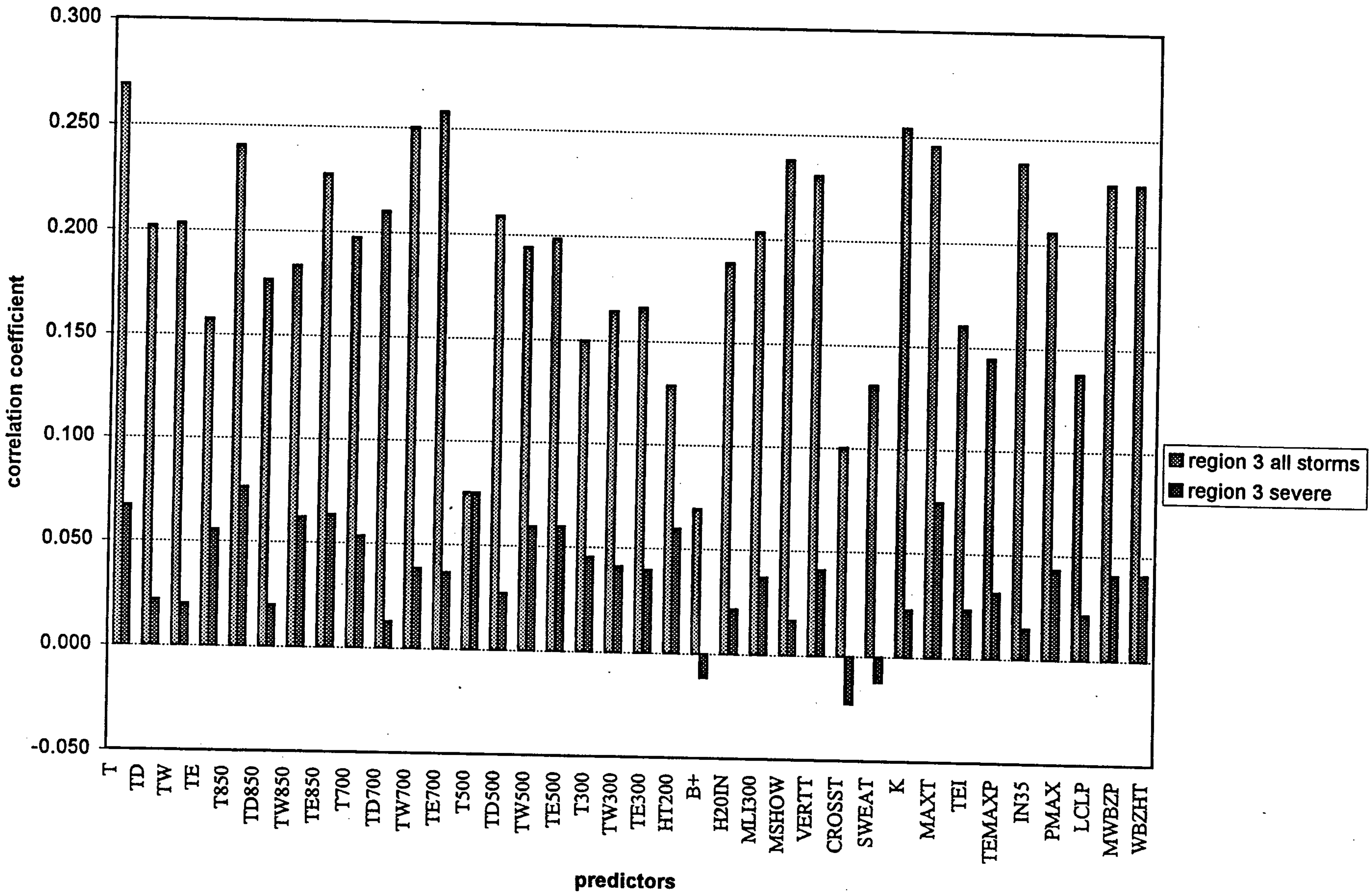


Figure 30. Same as Figure 28 but for 33 Statute Mile Radius.

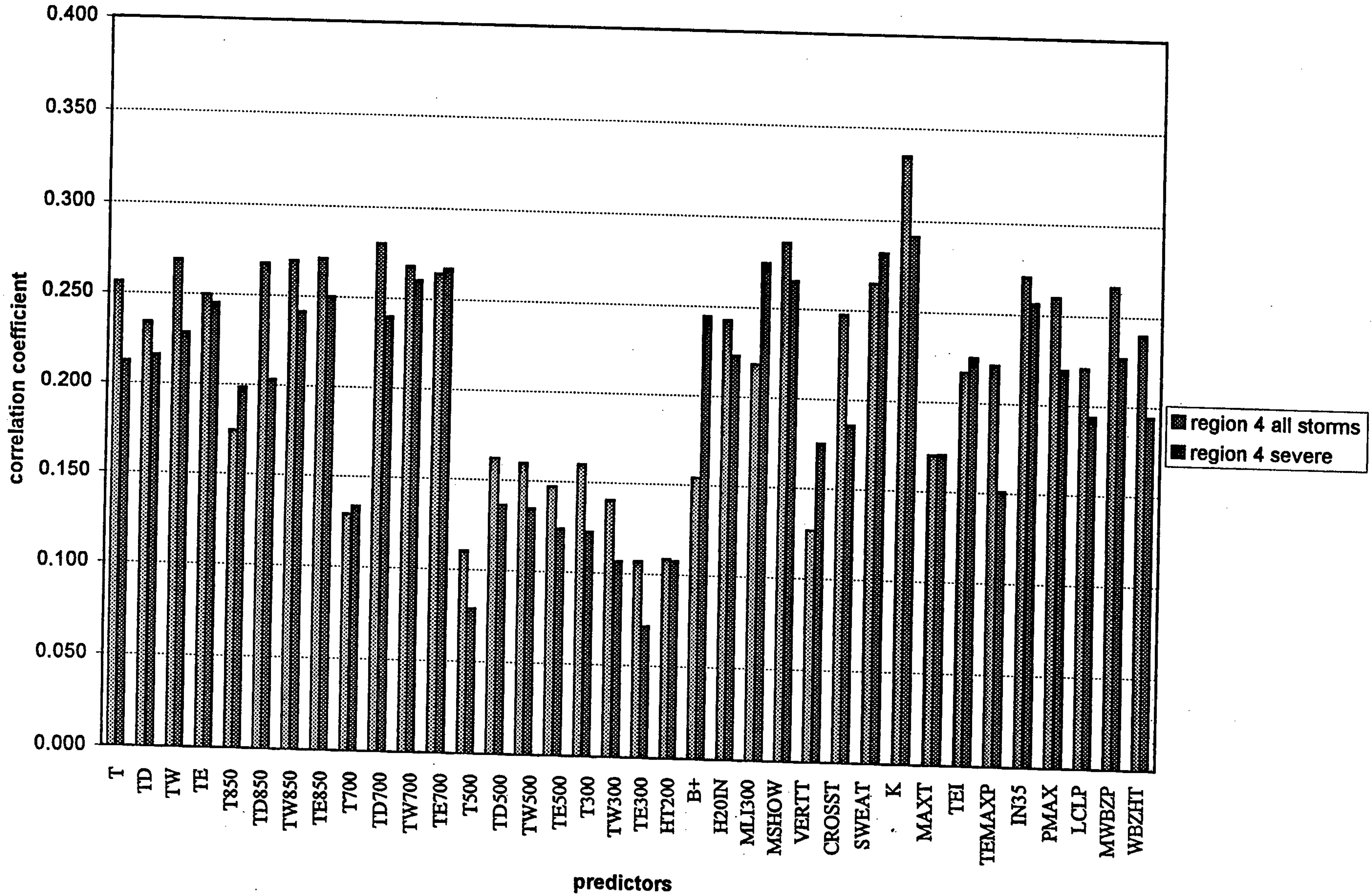


Figure 31. Correlation Coefficients for Region 4 (Panhandle) Within 100 Statute Mile Radius.

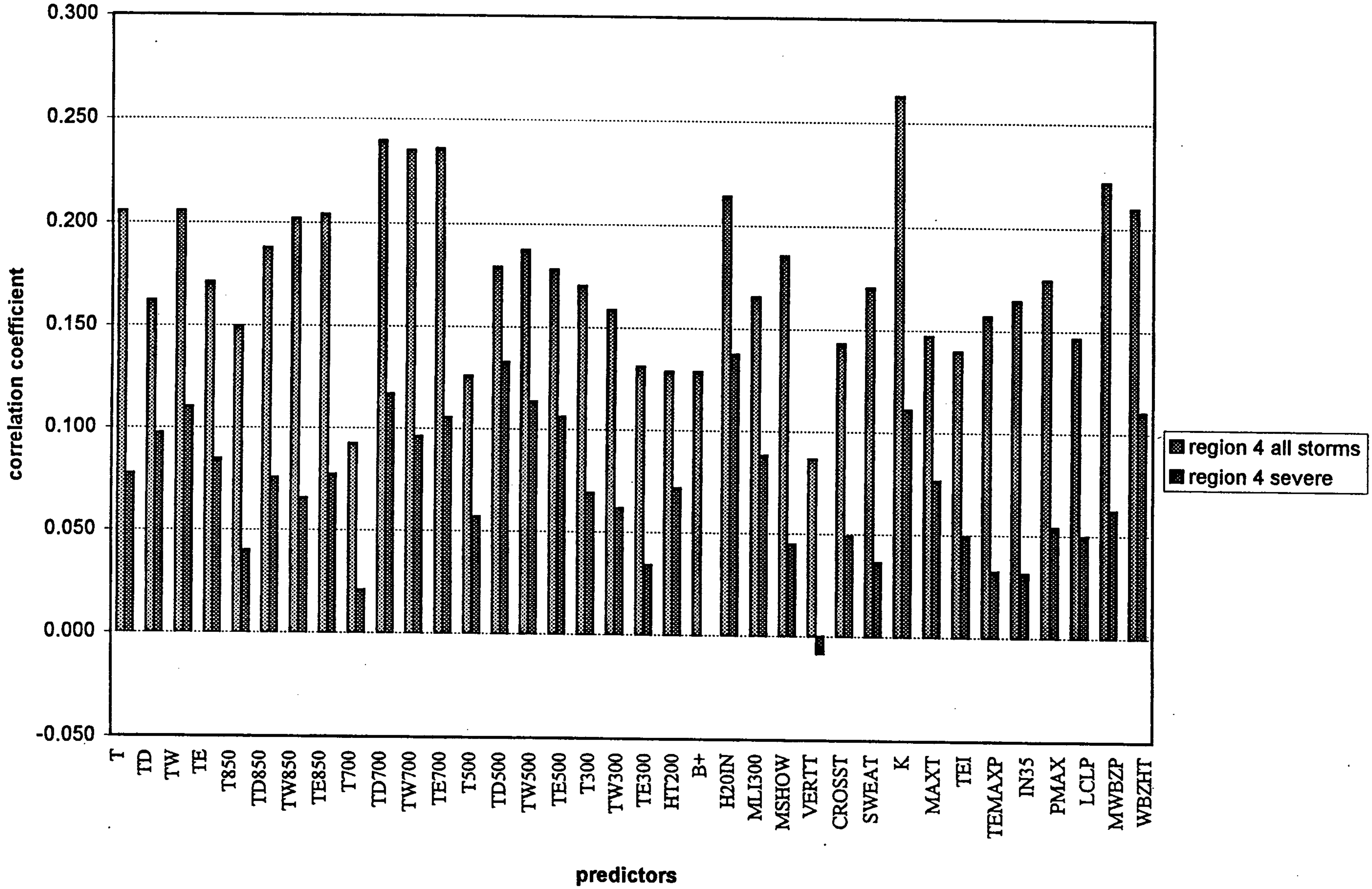


Figure 32. Same as Figure 31 but for 66 Statute Mile Radius.

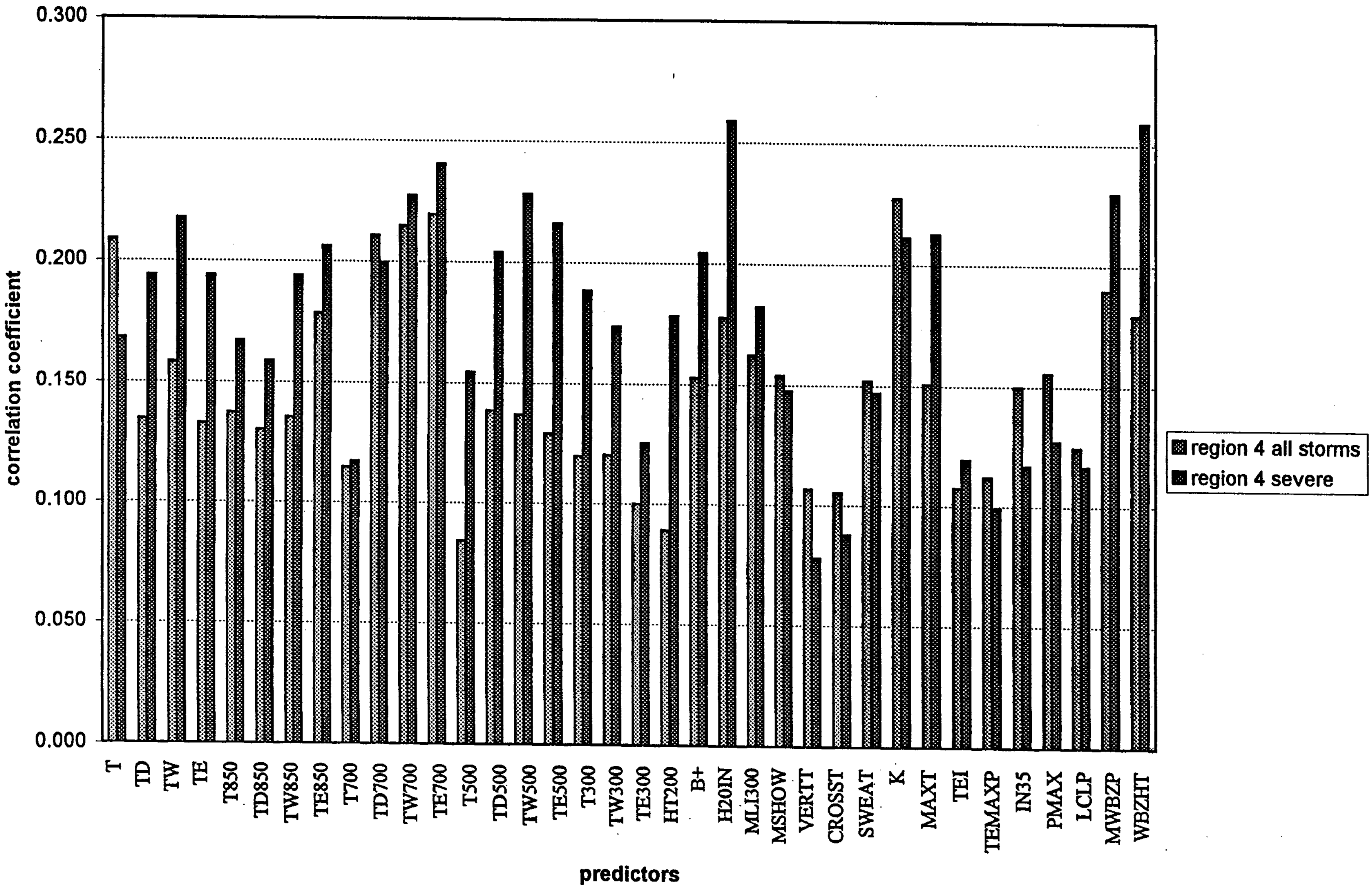


Figure 33. Same as Figure 31 but for 33 Statute Mile Radius.

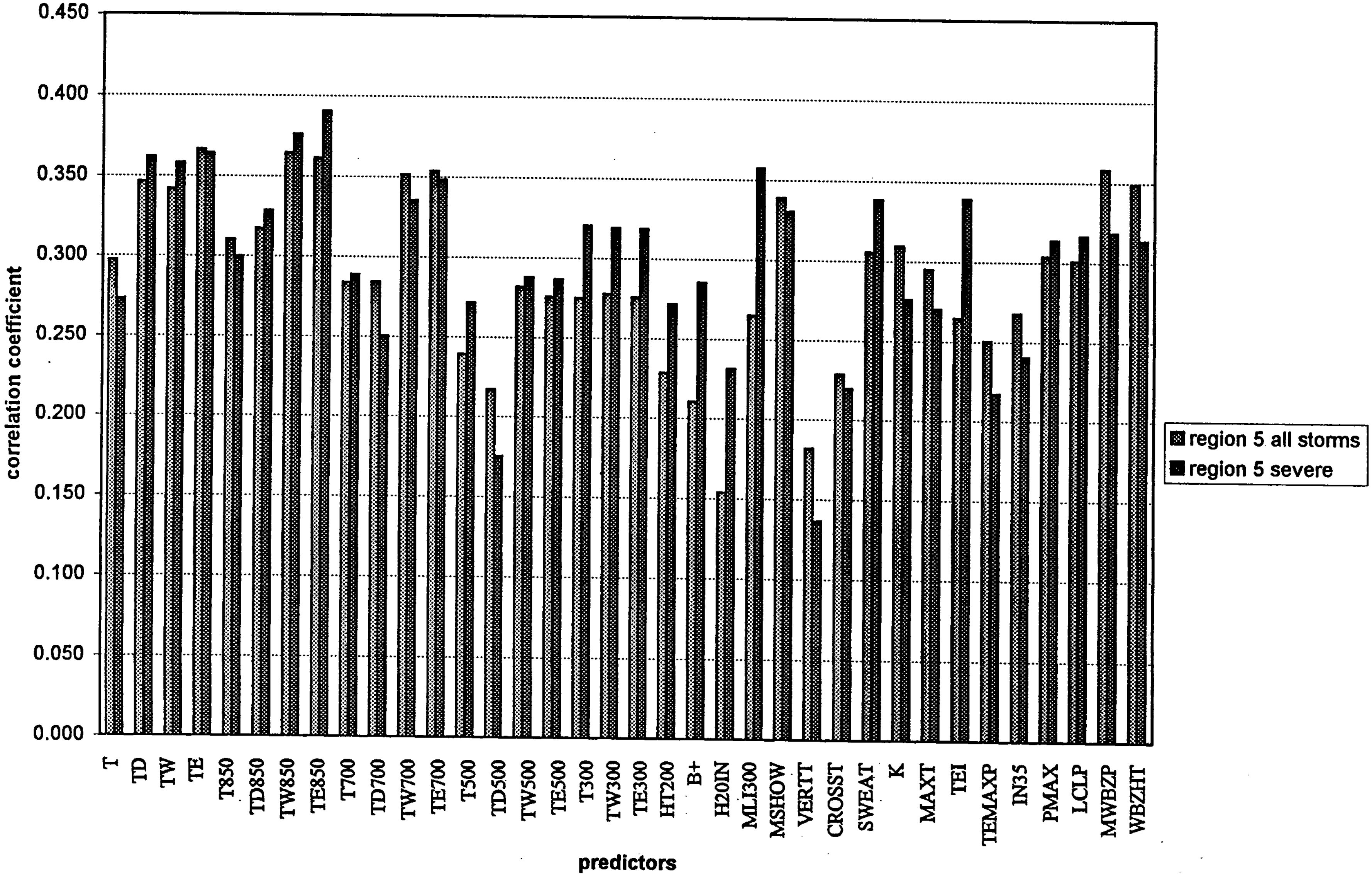


Figure 34. Correlation Coefficients for Region 5 (Midwest & Northern Tier) Within 100 Statute Mile Radius.

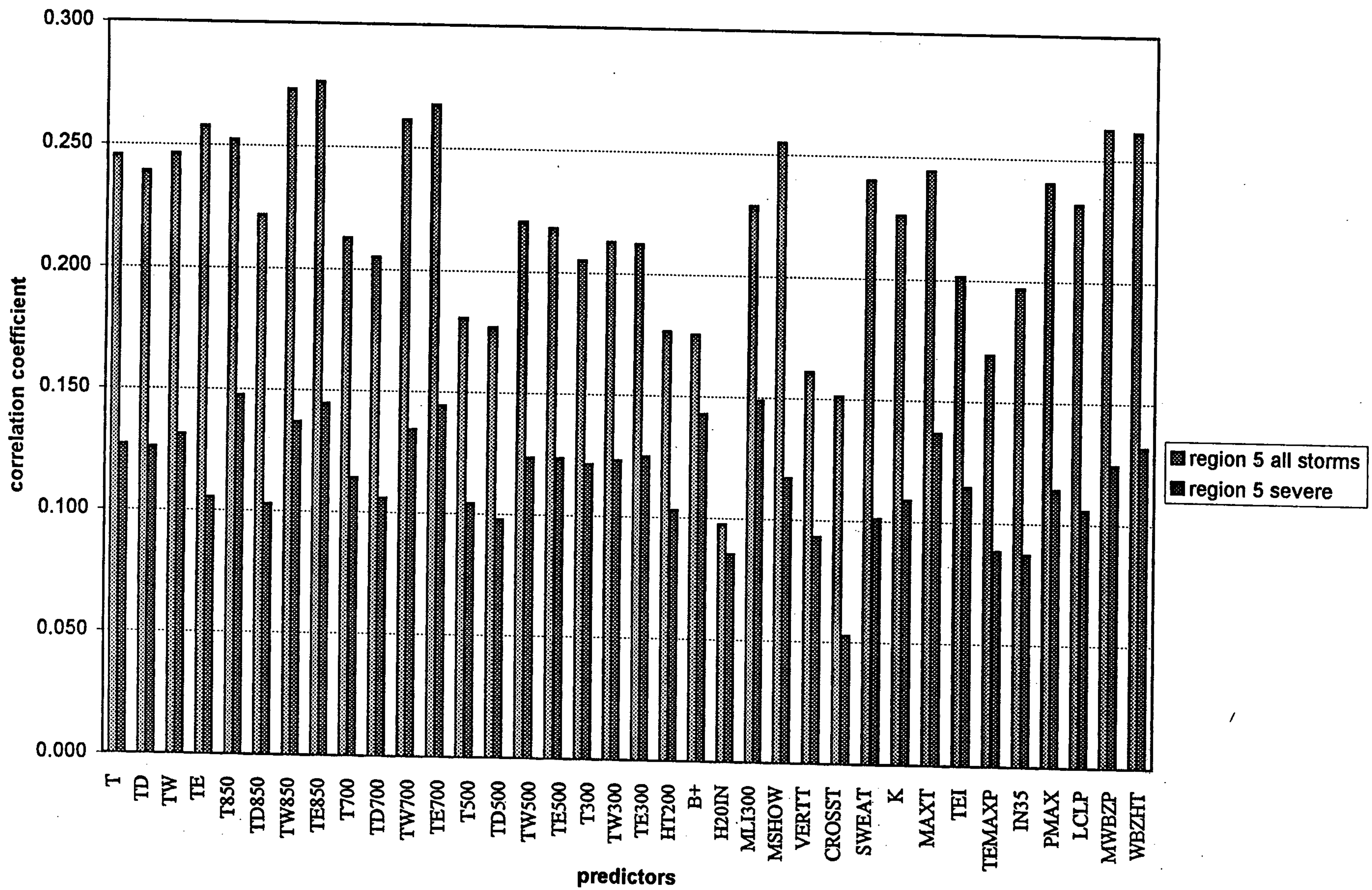


Figure 35. Same as Figure 34 but for 66 Statute Mile Radius.

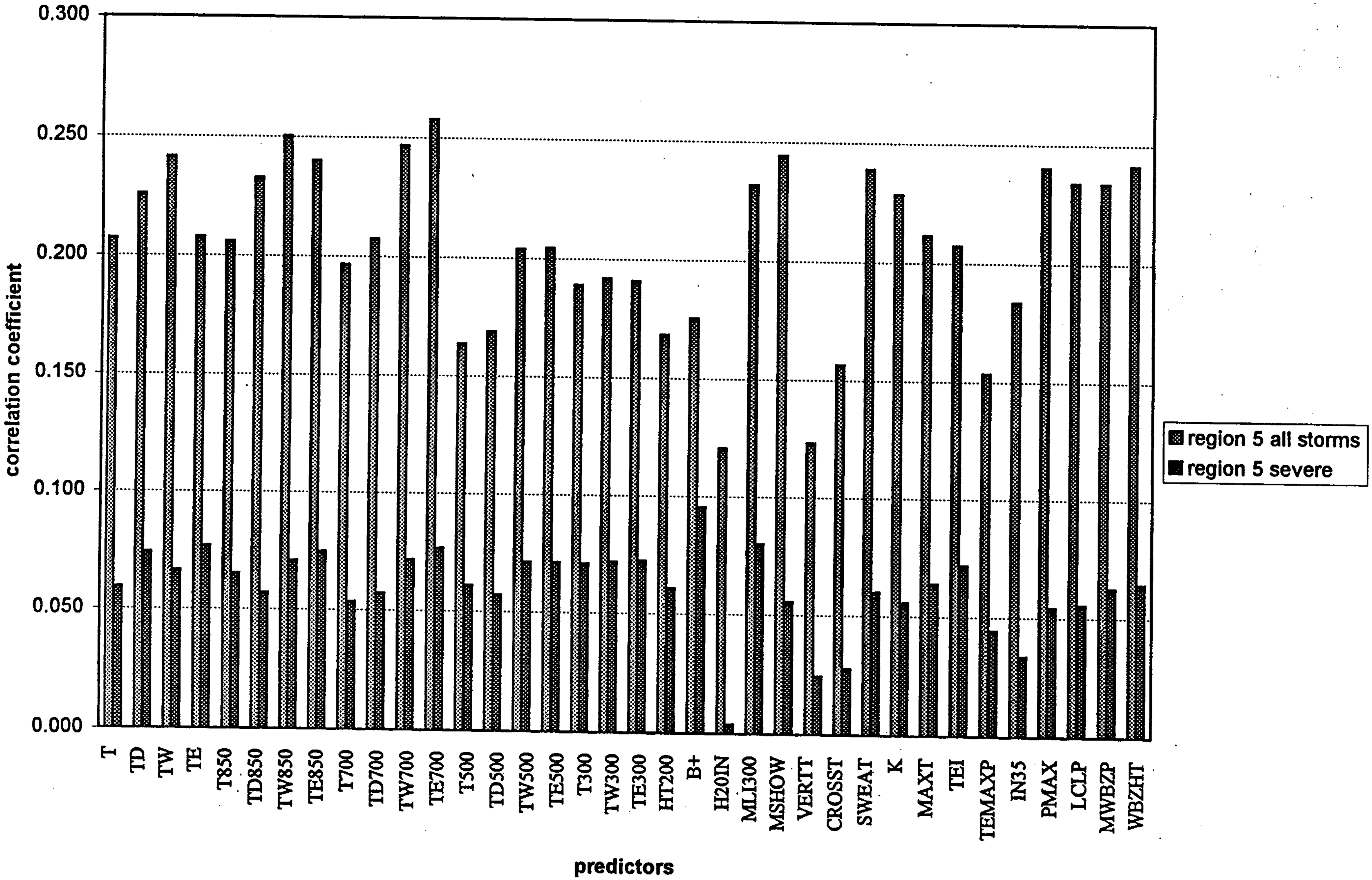


Figure 36. Same as Figure 34 but for 33 Statute Mile Radius.

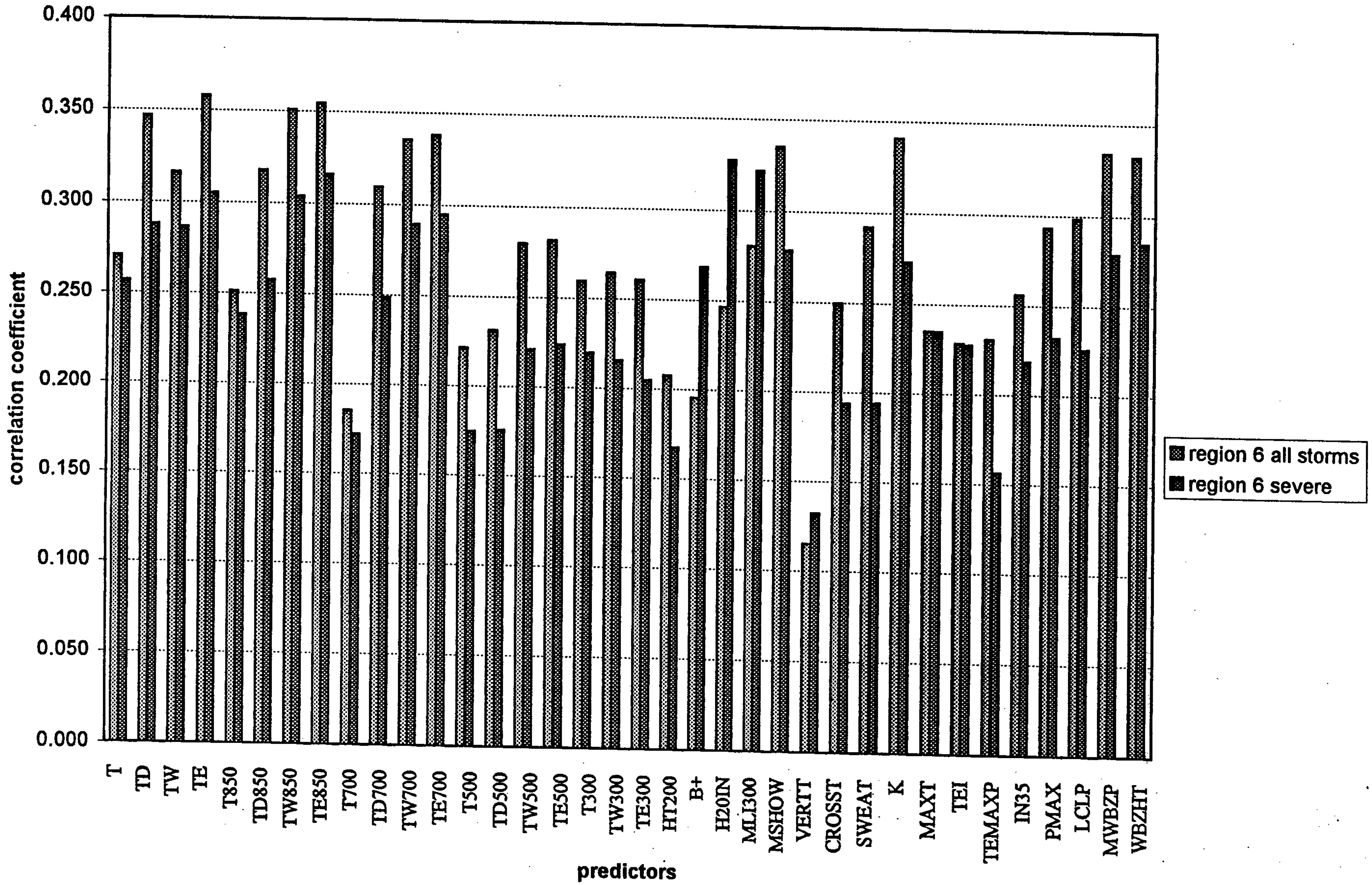


Figure 37. Correlation Coefficients for Region 6 (Southwest Interior) Within 100 Statute Mile Radius.

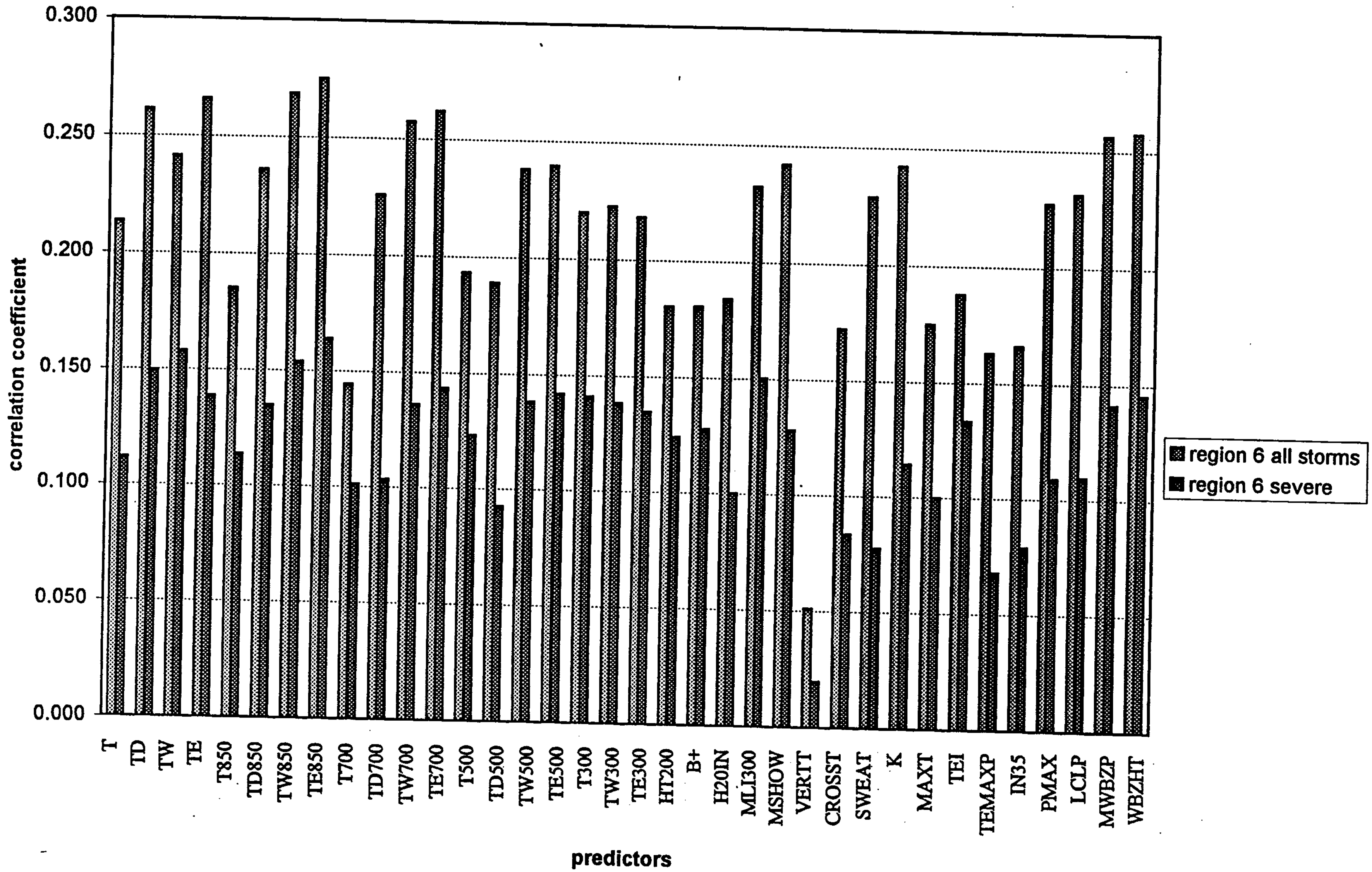


Figure 38. Same as Figure 37 but for 66 Statute Mile Radius.

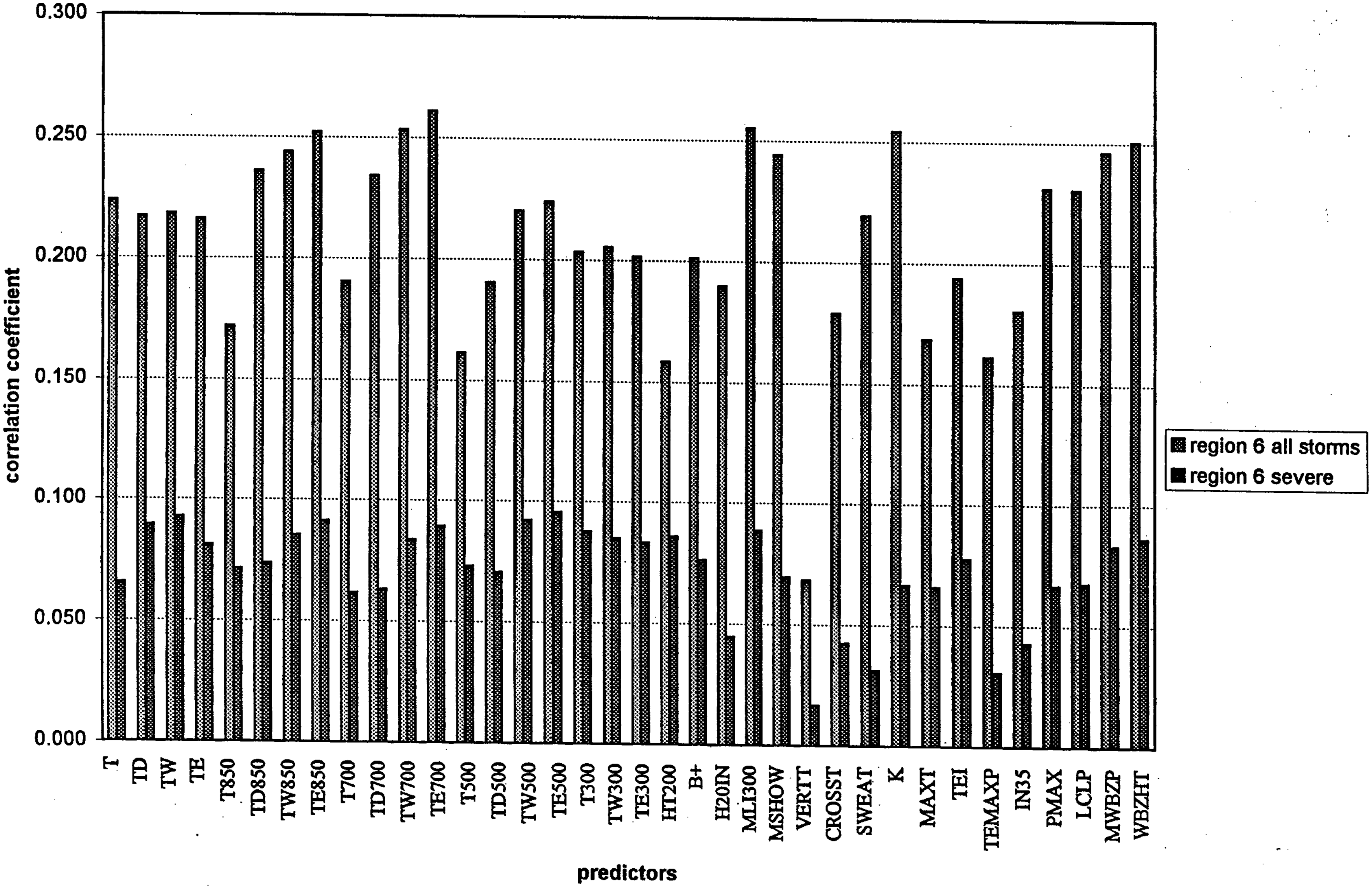


Figure 39. Same as Figure 37 but for 33 Statute Mile Radius.

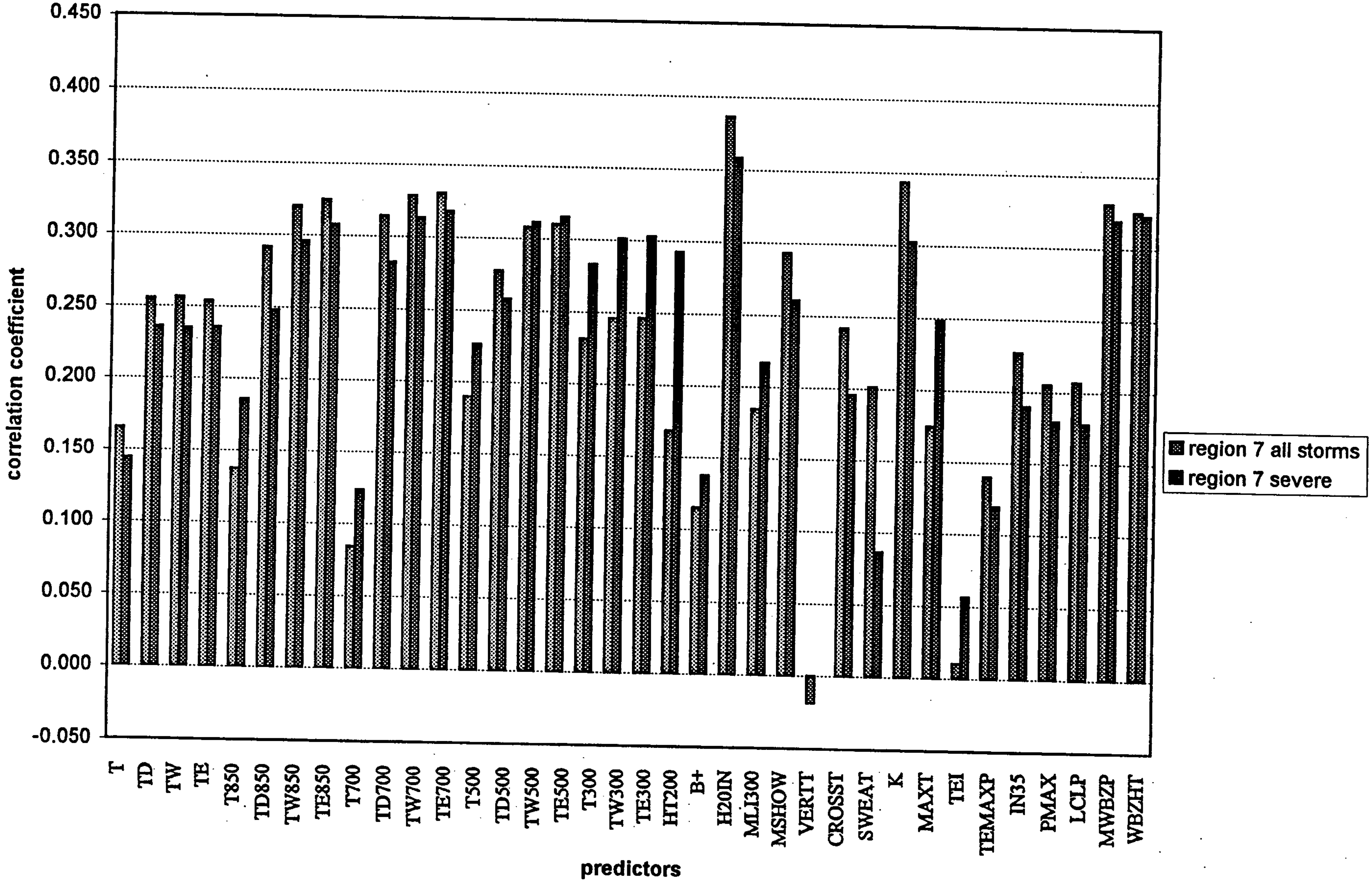


Figure 40. Correlation Coefficients for Region 7 (Gulf Coast) Within 100 Statute Mile Radius.

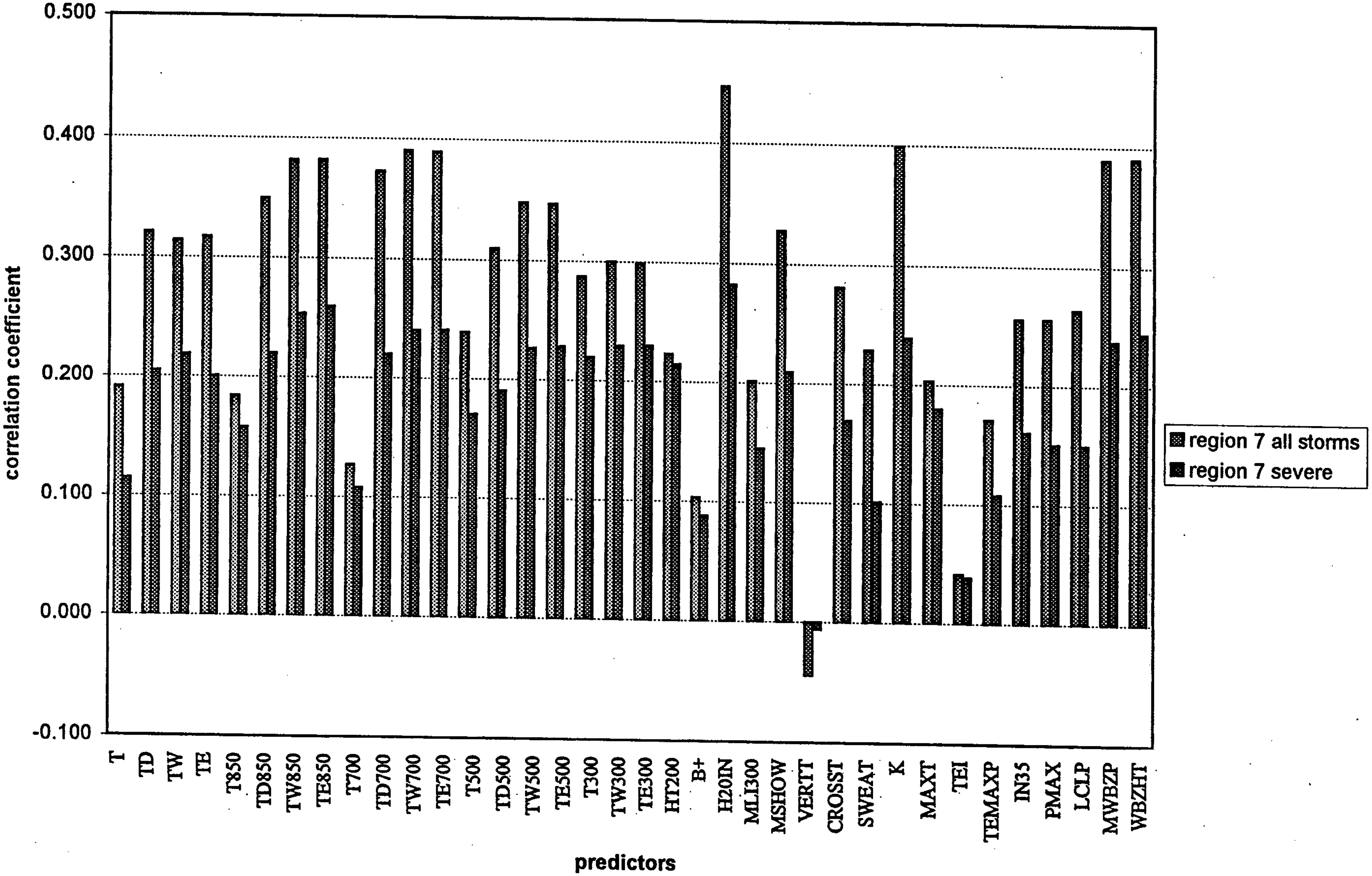


Figure 41. Same as Figure 40 but for 66 Statute Mile Radius.

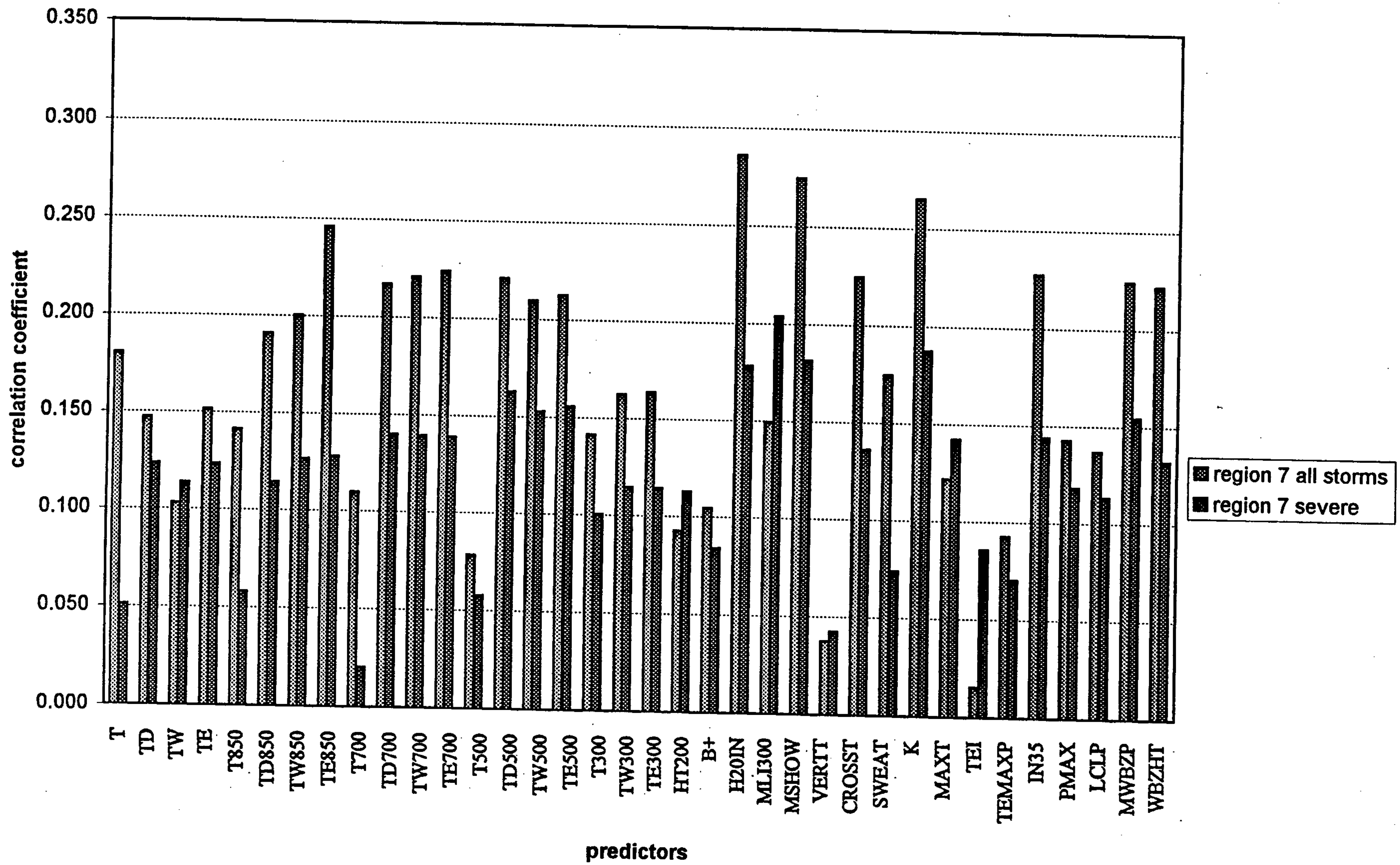


Figure 42. Same as Figure 40 but for 33 Statute Mile Radius.

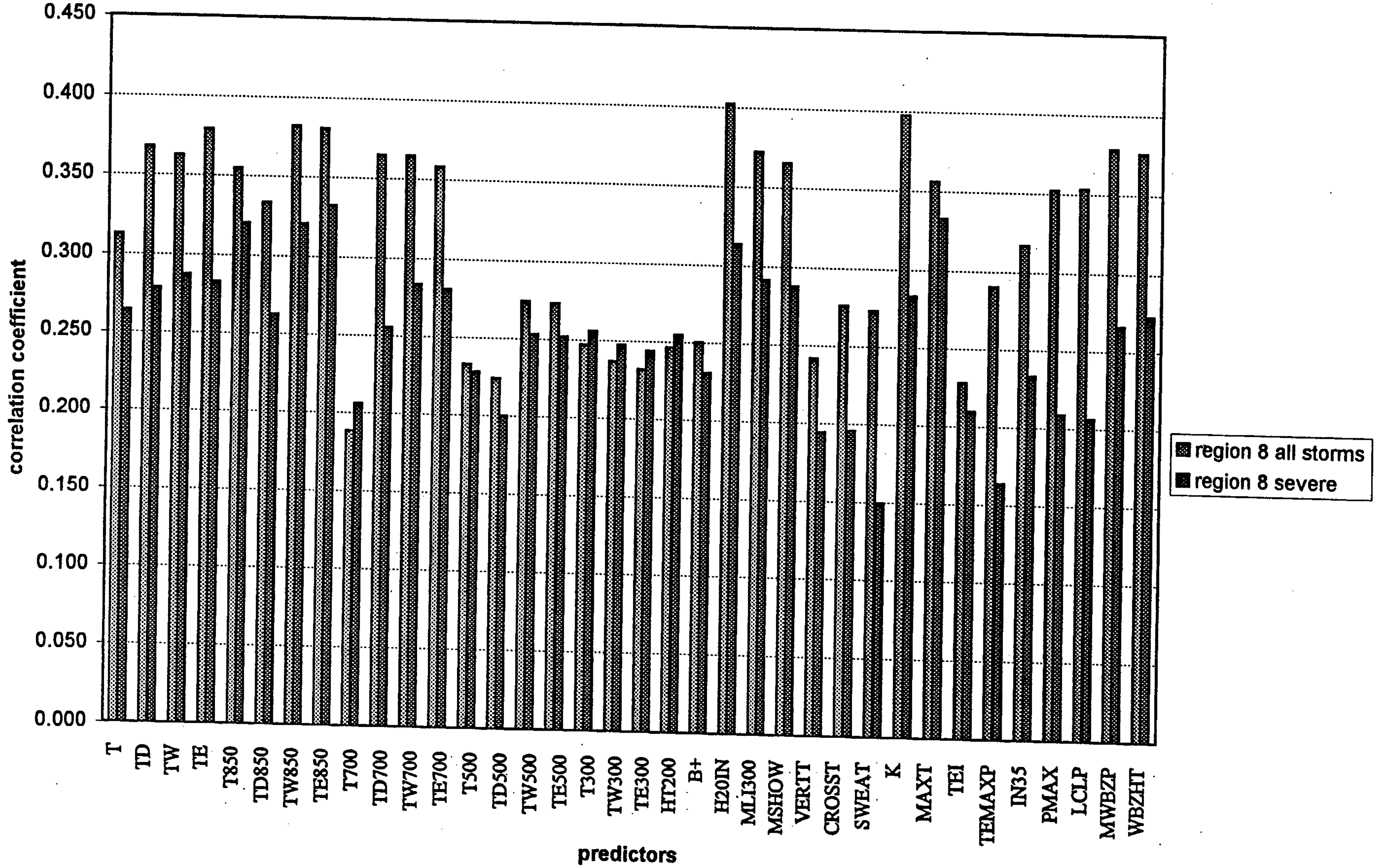


Figure 43. Correlation Coefficients for Region 8 (Southeast Coast) Within 100 Statute Mile Radius.

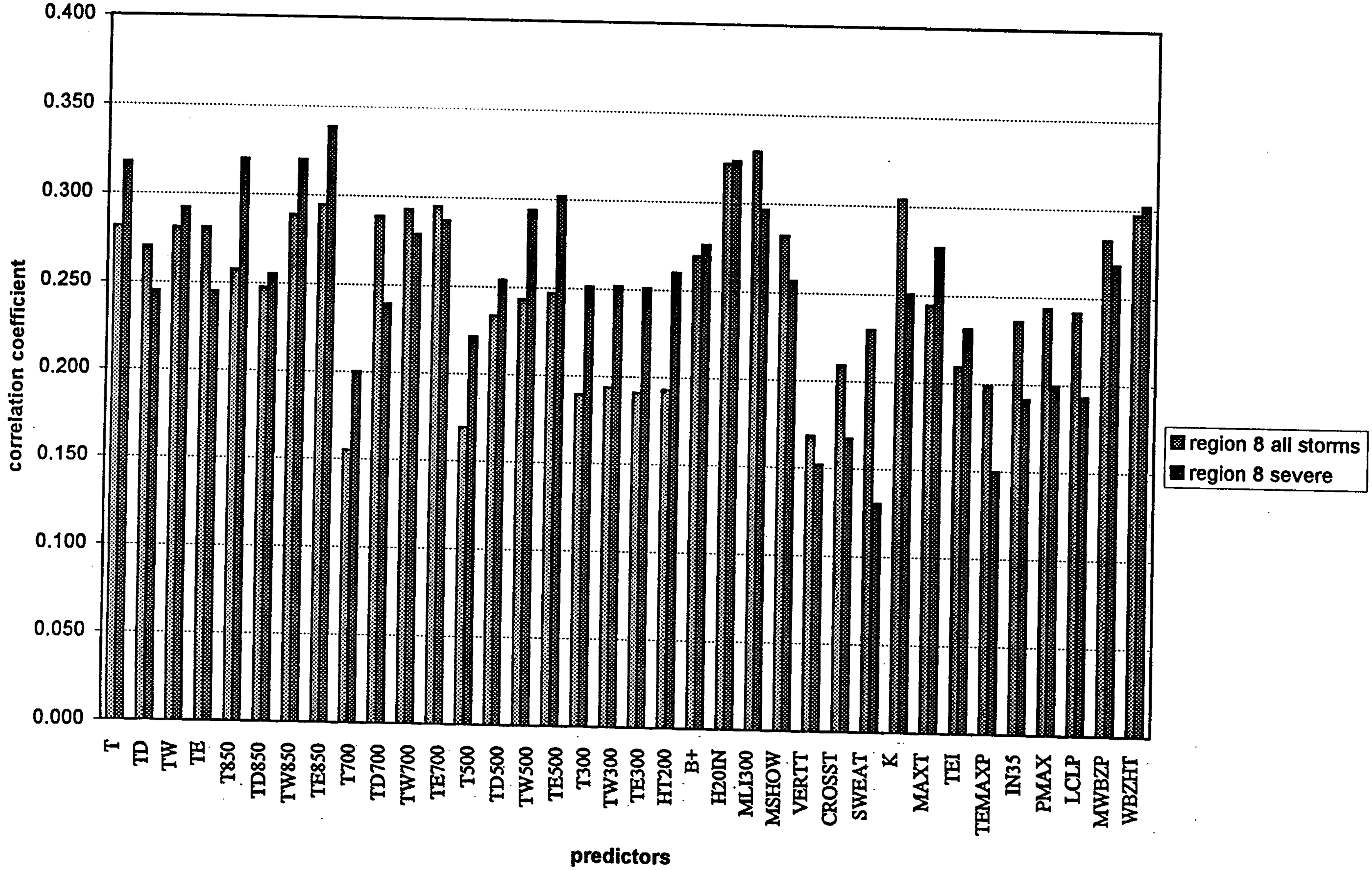


Figure 44. Same as Figure 43 but for 66 Statute Mile Radius.

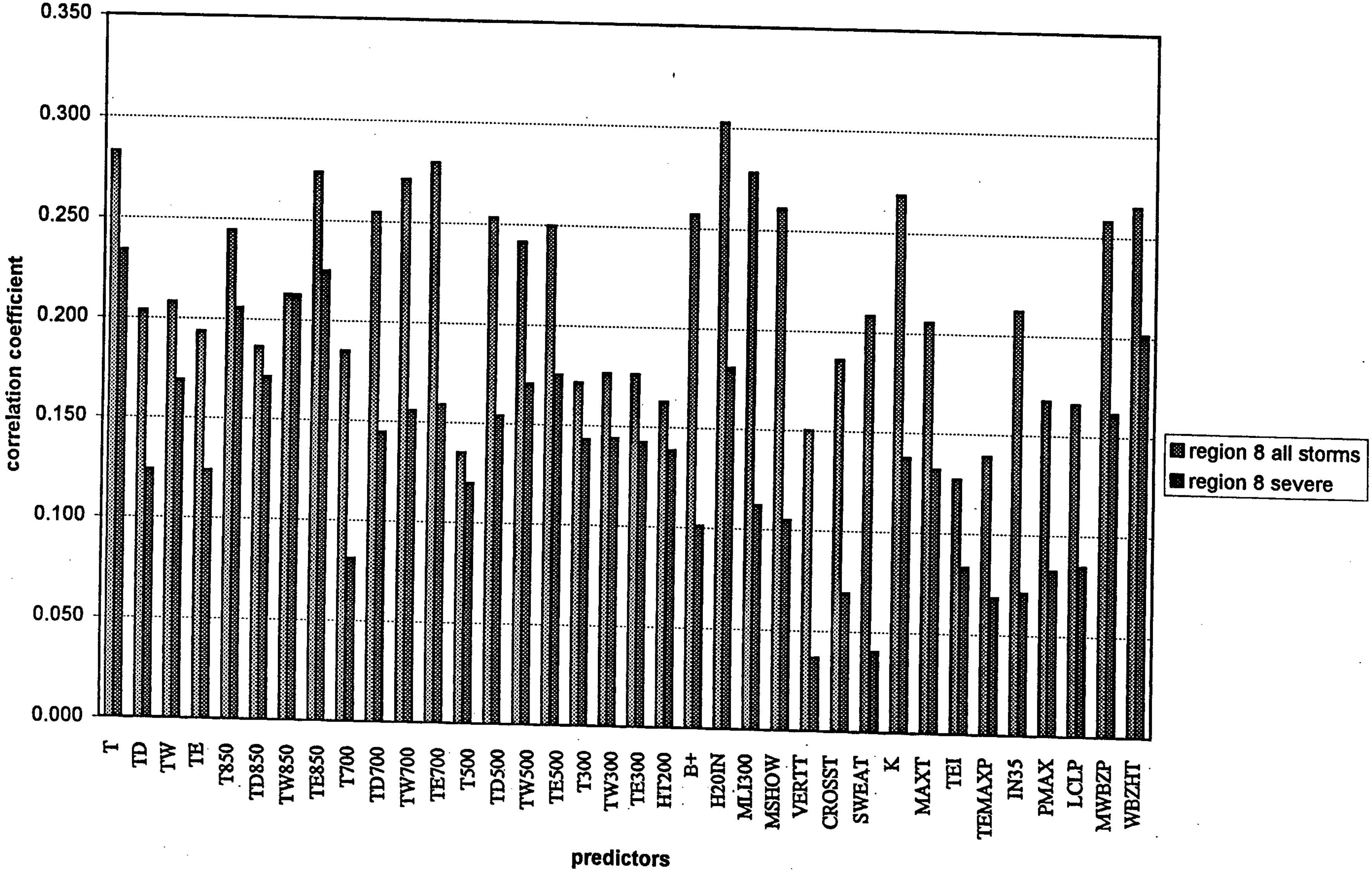


Figure 45. Same as Figure 43 but for 33 Statute Mile Radius.

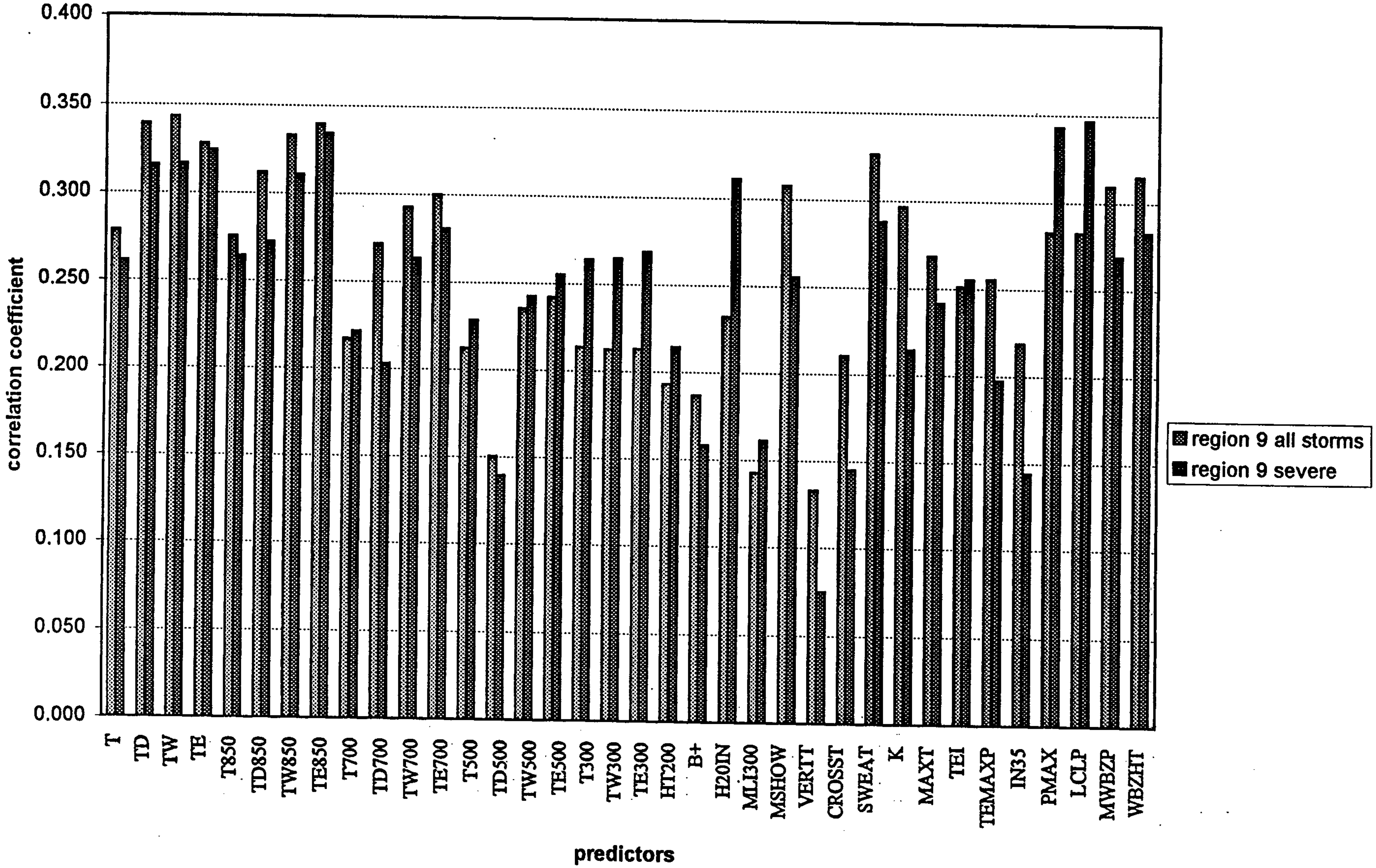


Figure 46. Correlation Coefficients for Region 9 (Great Lakes) Within 100 Statute Mile Radius.

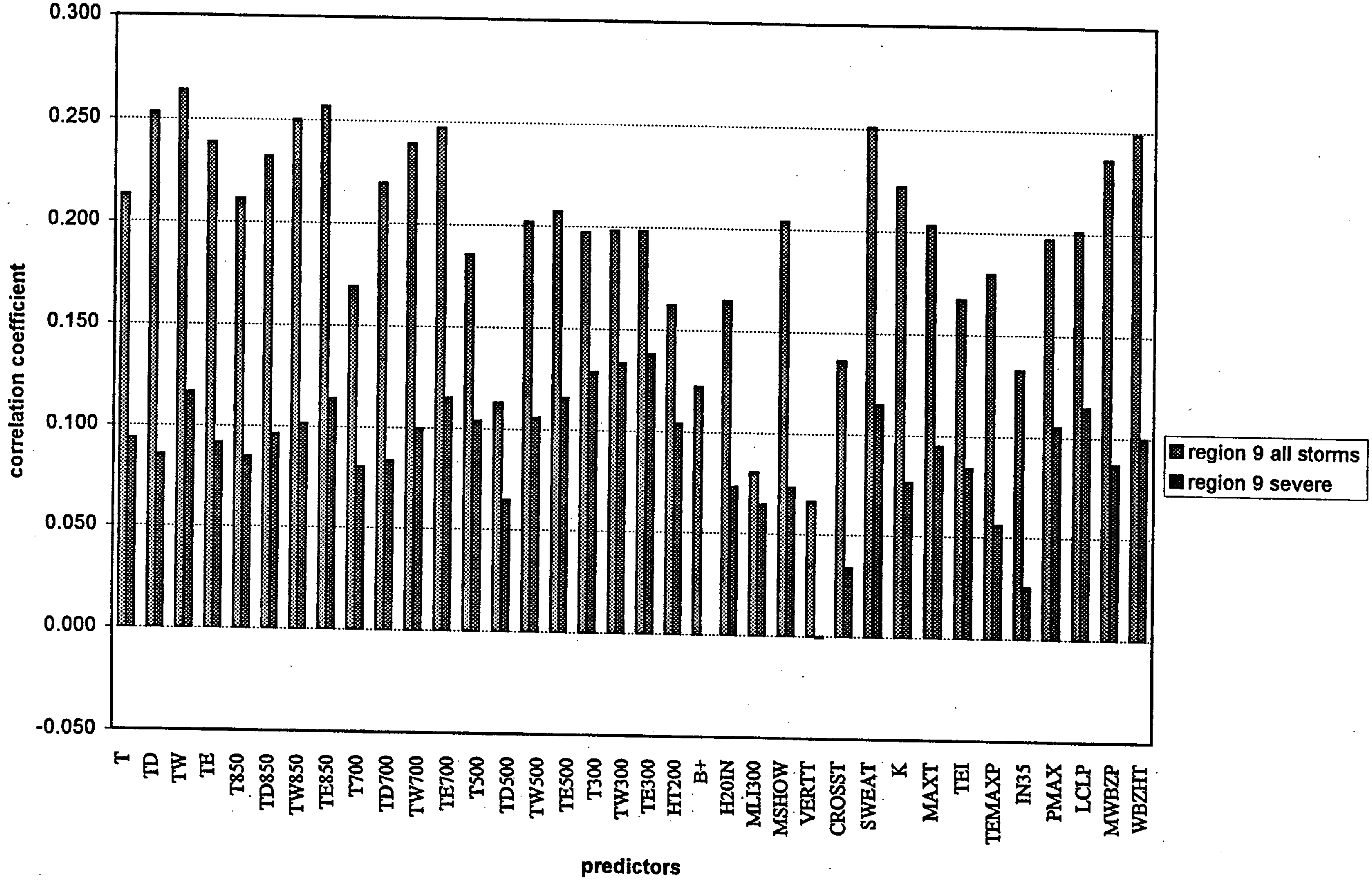


Figure 47. Same as Figure 46 but for 66 Statute Mile Radius.

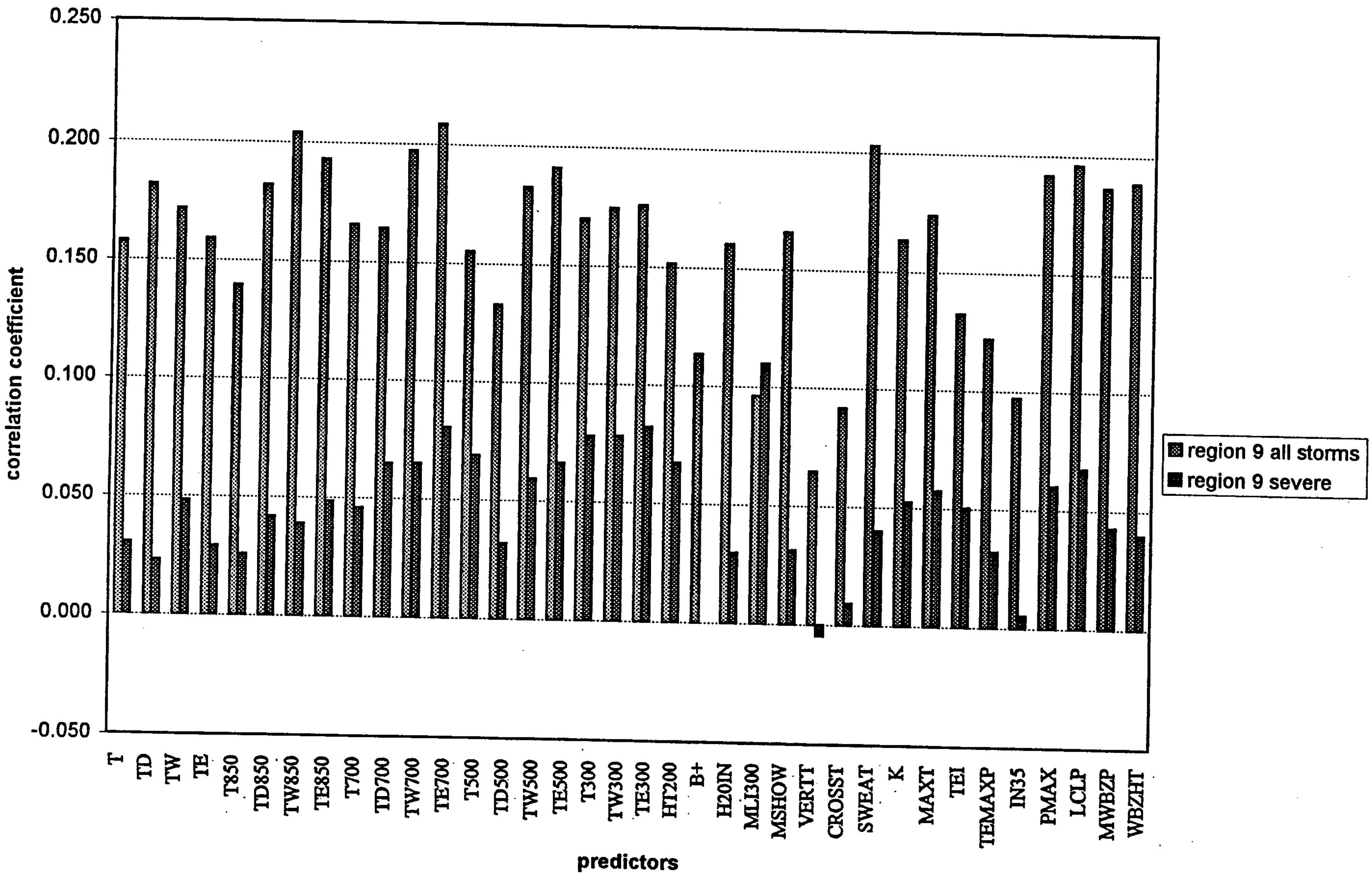


Figure 48. Same as Figure 46 but for 33 Statute Mile Radius.

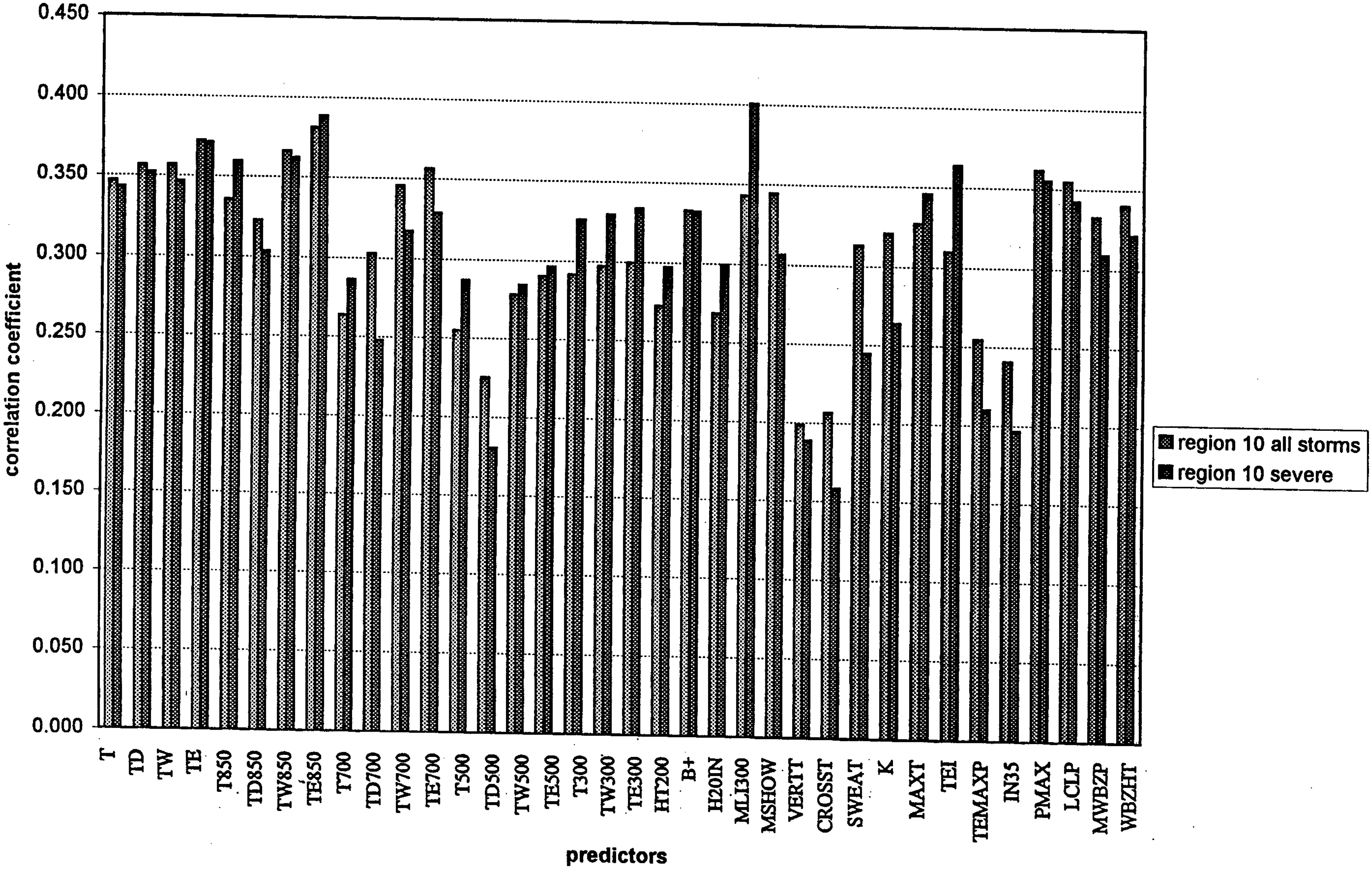


Figure 49. Correlation Coefficients for Region 10 (Appalachians) Within 100 Statute Mile Radius.

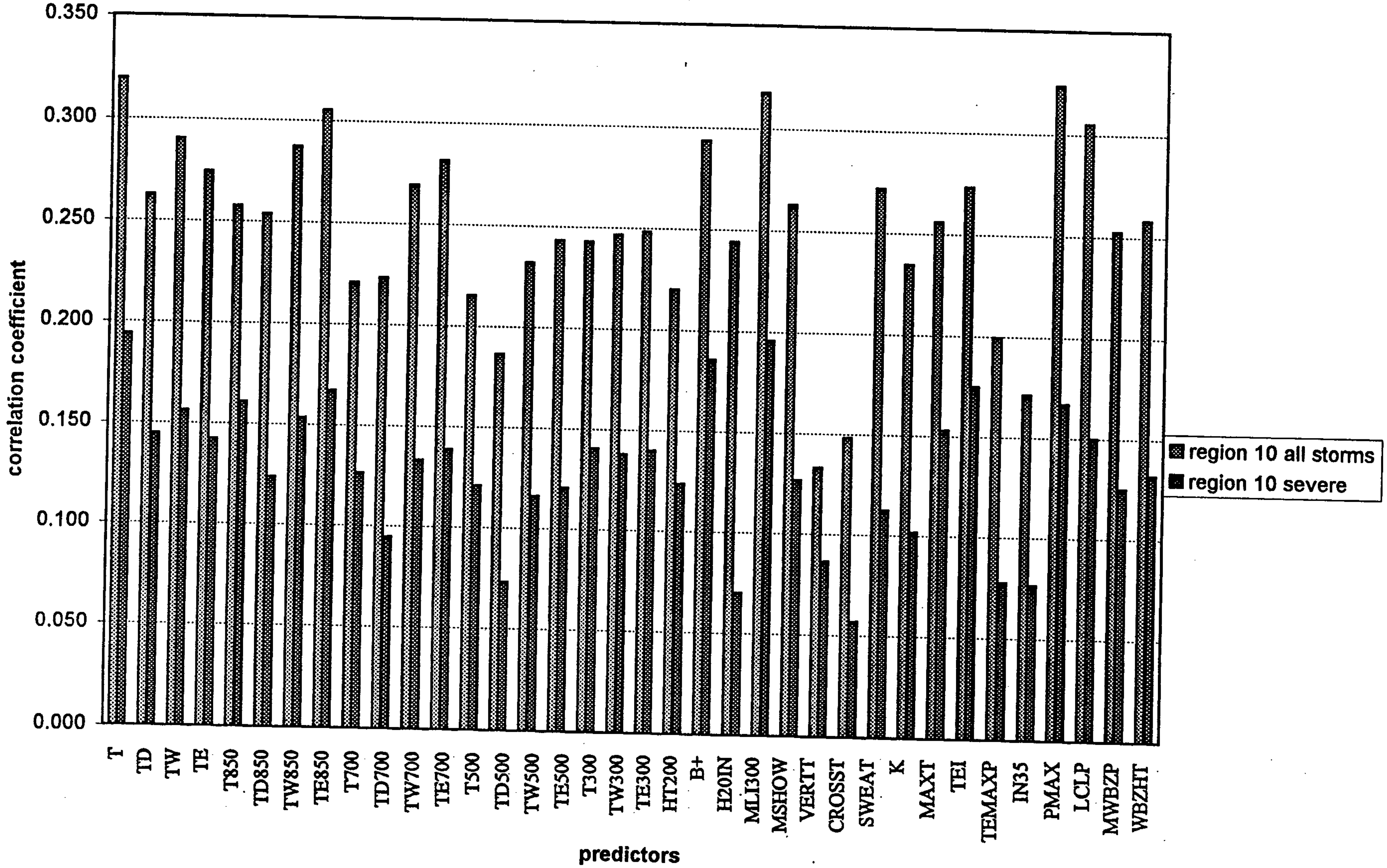


Figure 50. Same as Figure 49 but for 66 Statute Mile Radius.

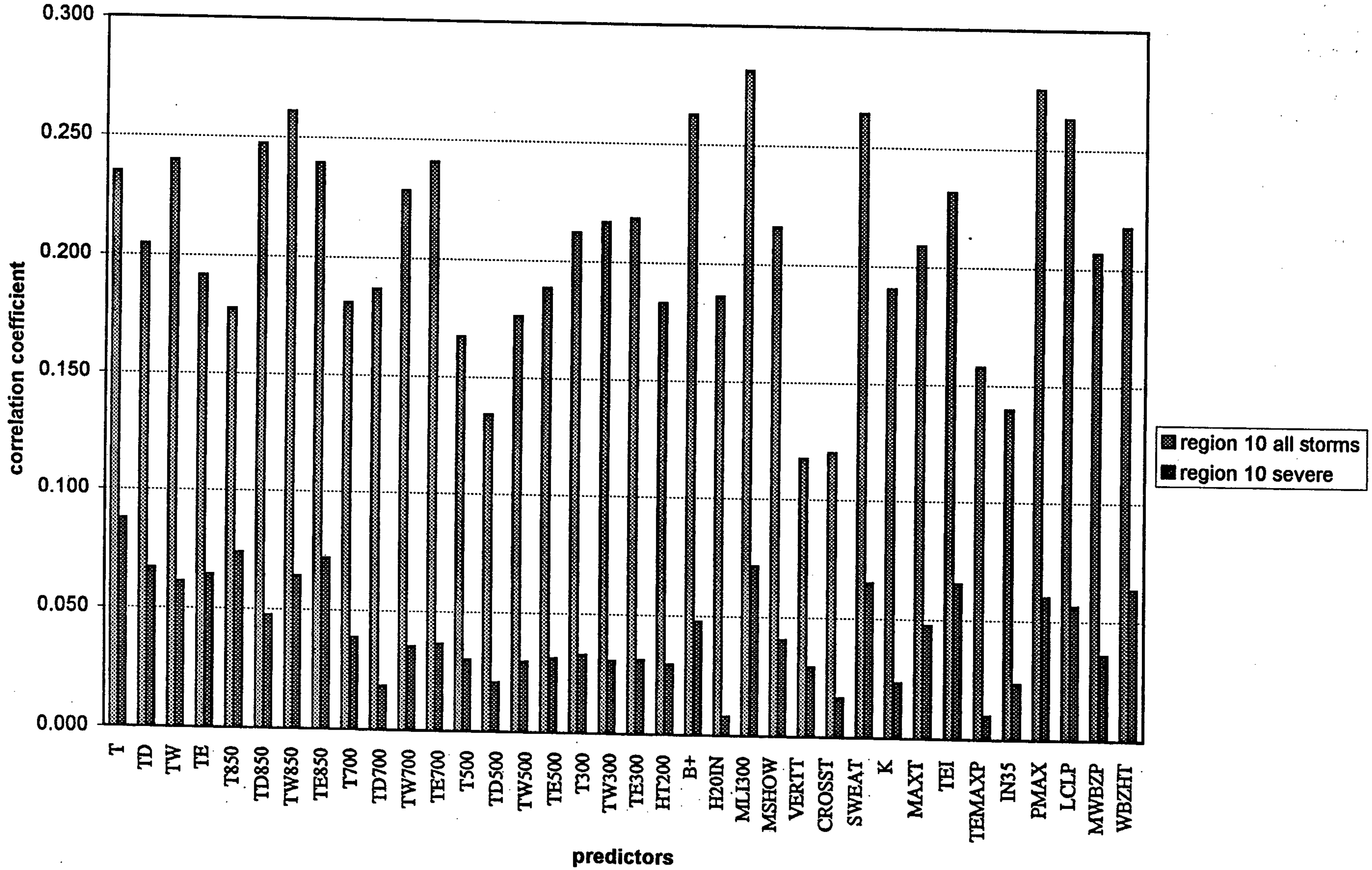


Figure 51. Same as Figure 49 but for 33 Statute Mile Radius.

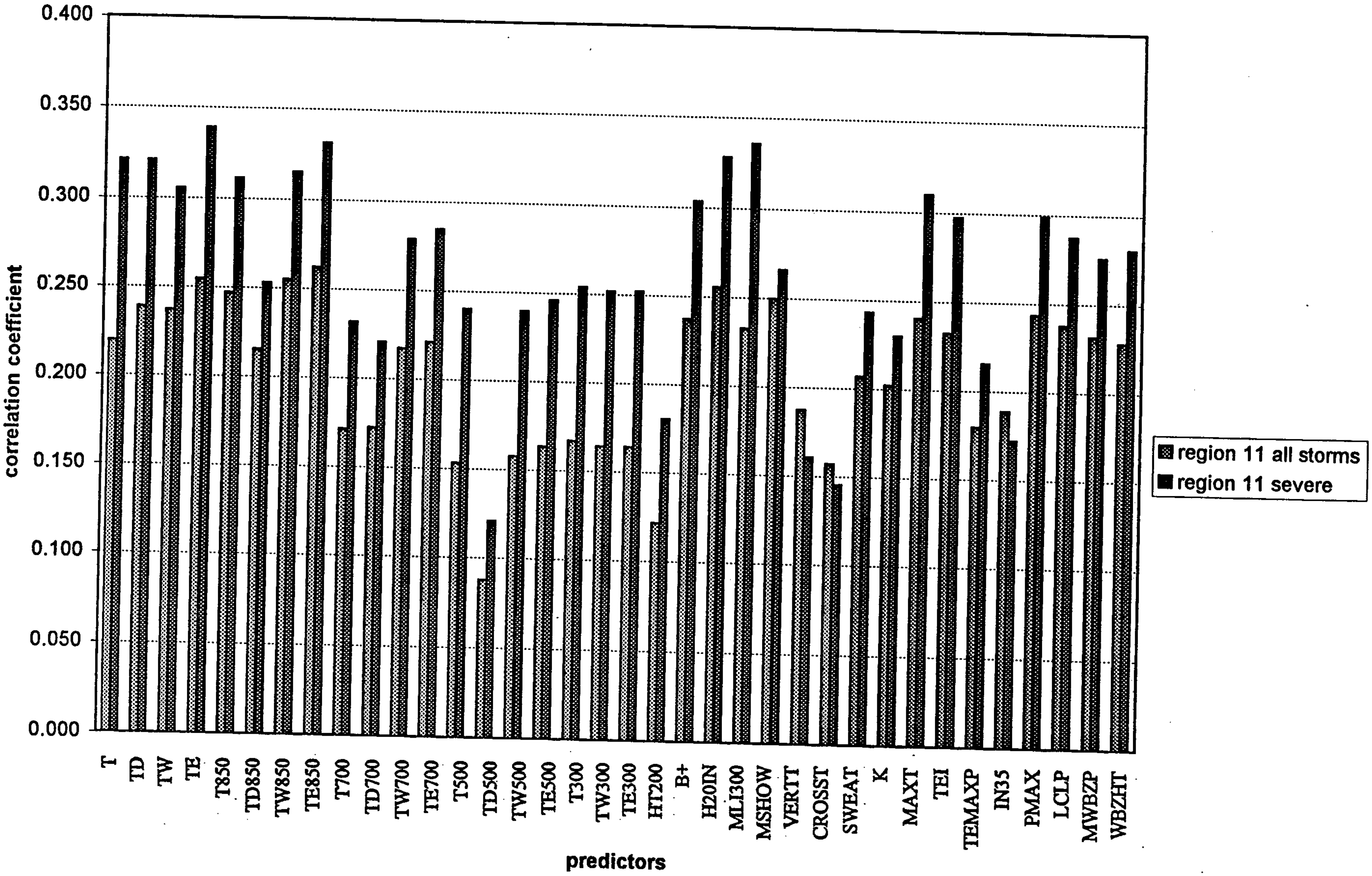


Figure 52. Correlation Coefficients for Region 11 (Northeast Coast) Within 100 Statute Mile radius.

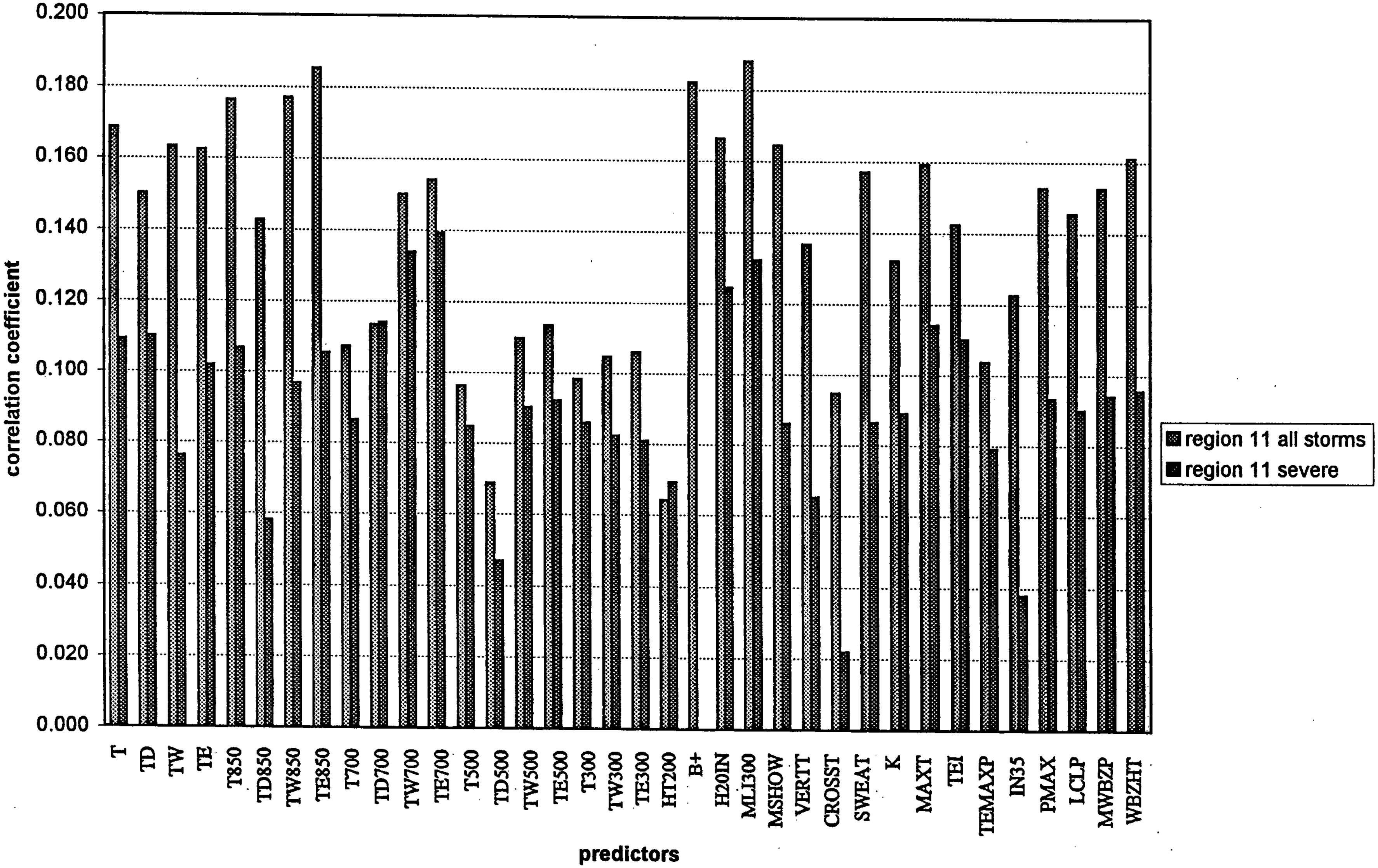


Figure 53. Same as Figure 52 but for 66 Statute Mile Radius.

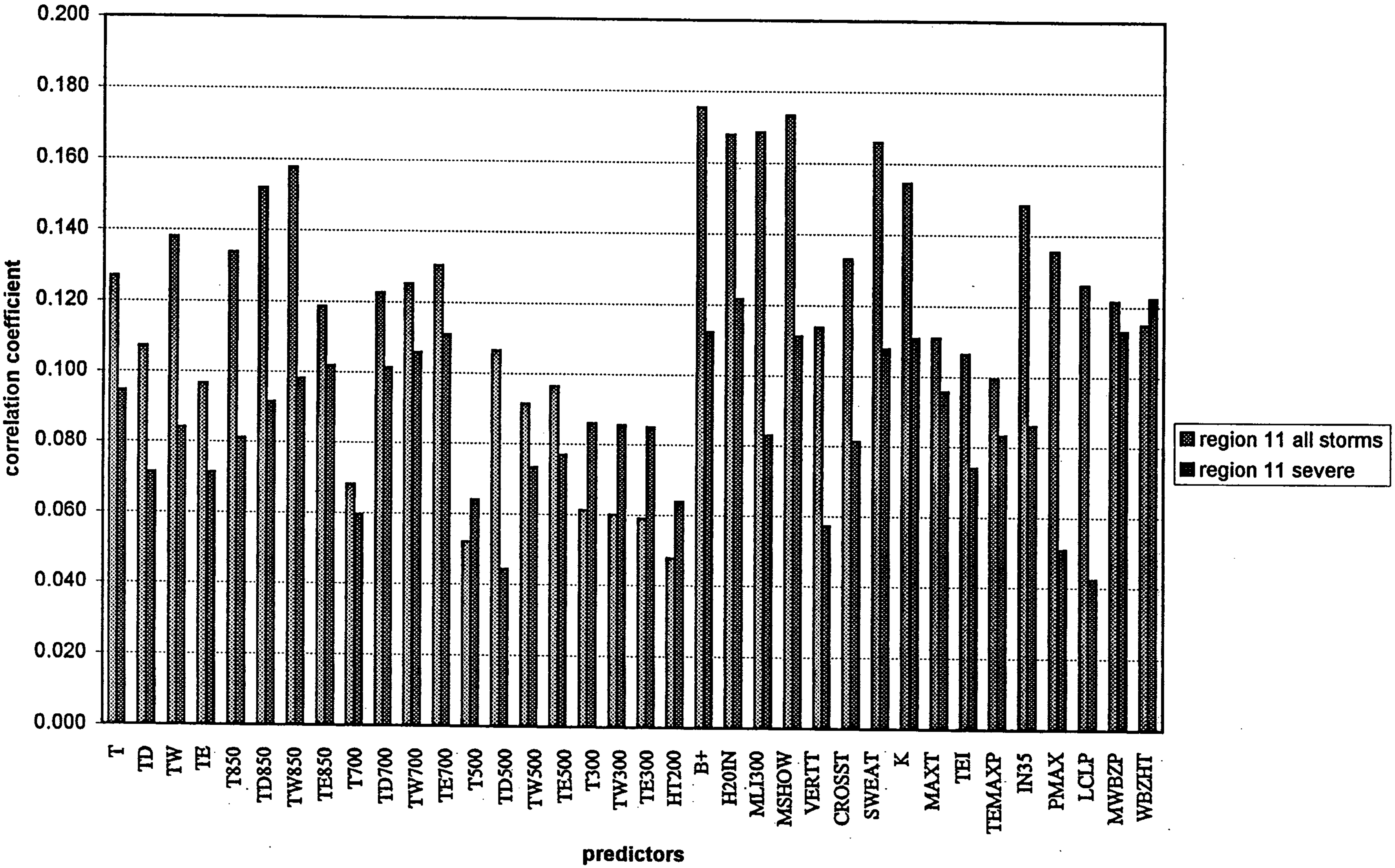


Figure 54. Same as Figure 52 but for 33 Statute Mile Radius.

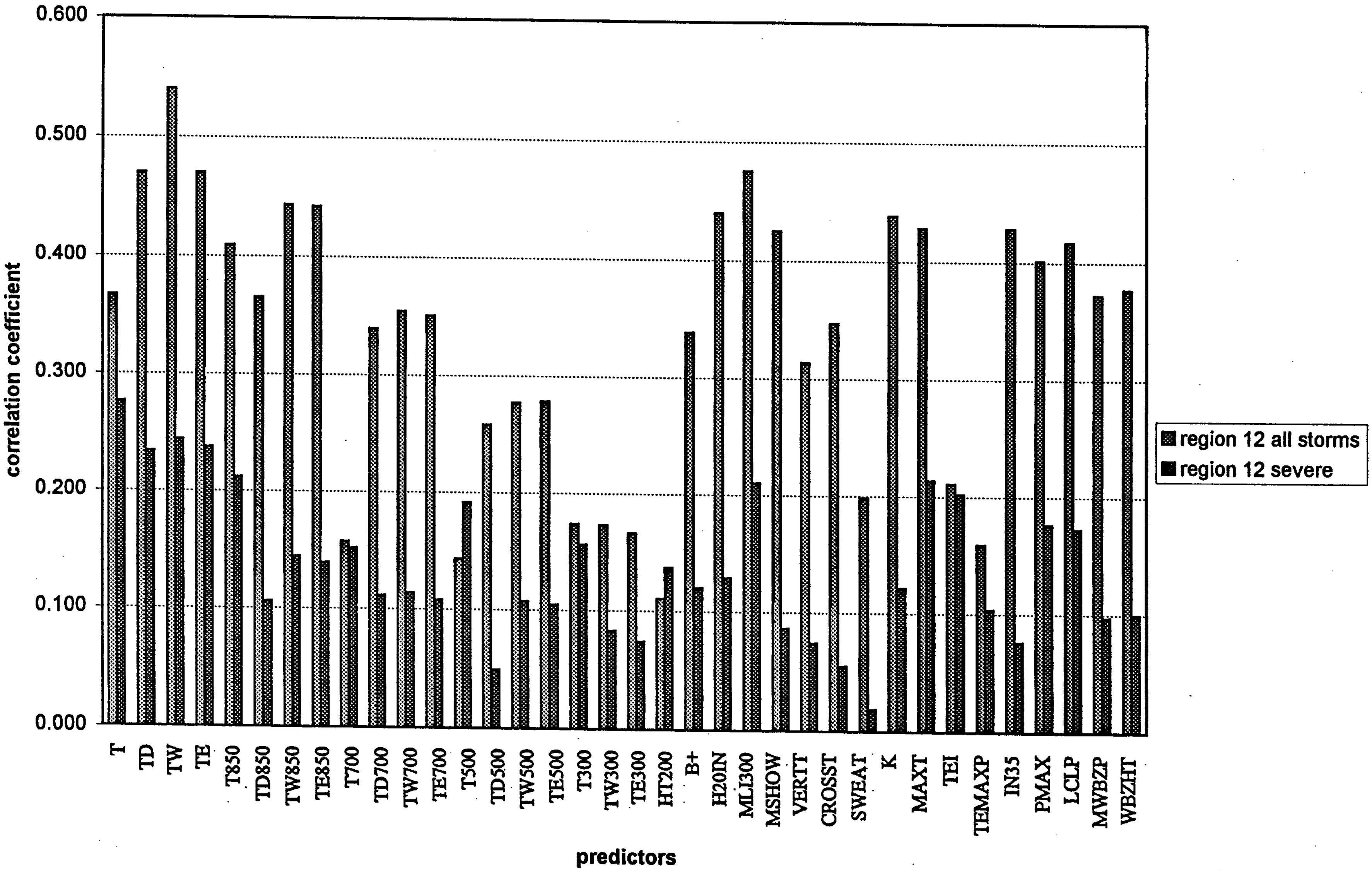


Figure 55. Correlation Coefficients for Region 12 (Southeast) Within 100 Statute Mile Radius.

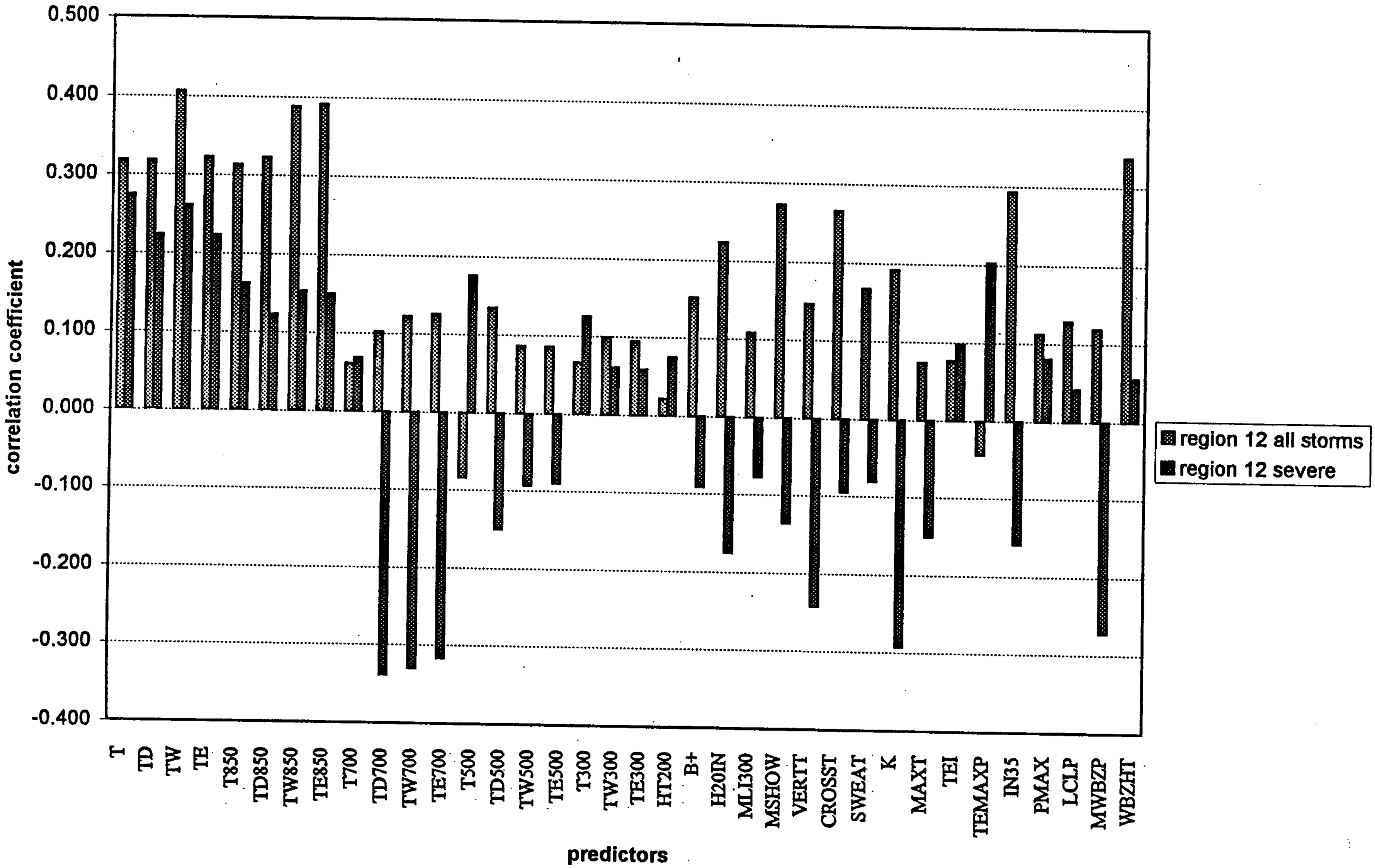


Figure 56. Same as Figure 55 but for 66 Statute Mile Radius.

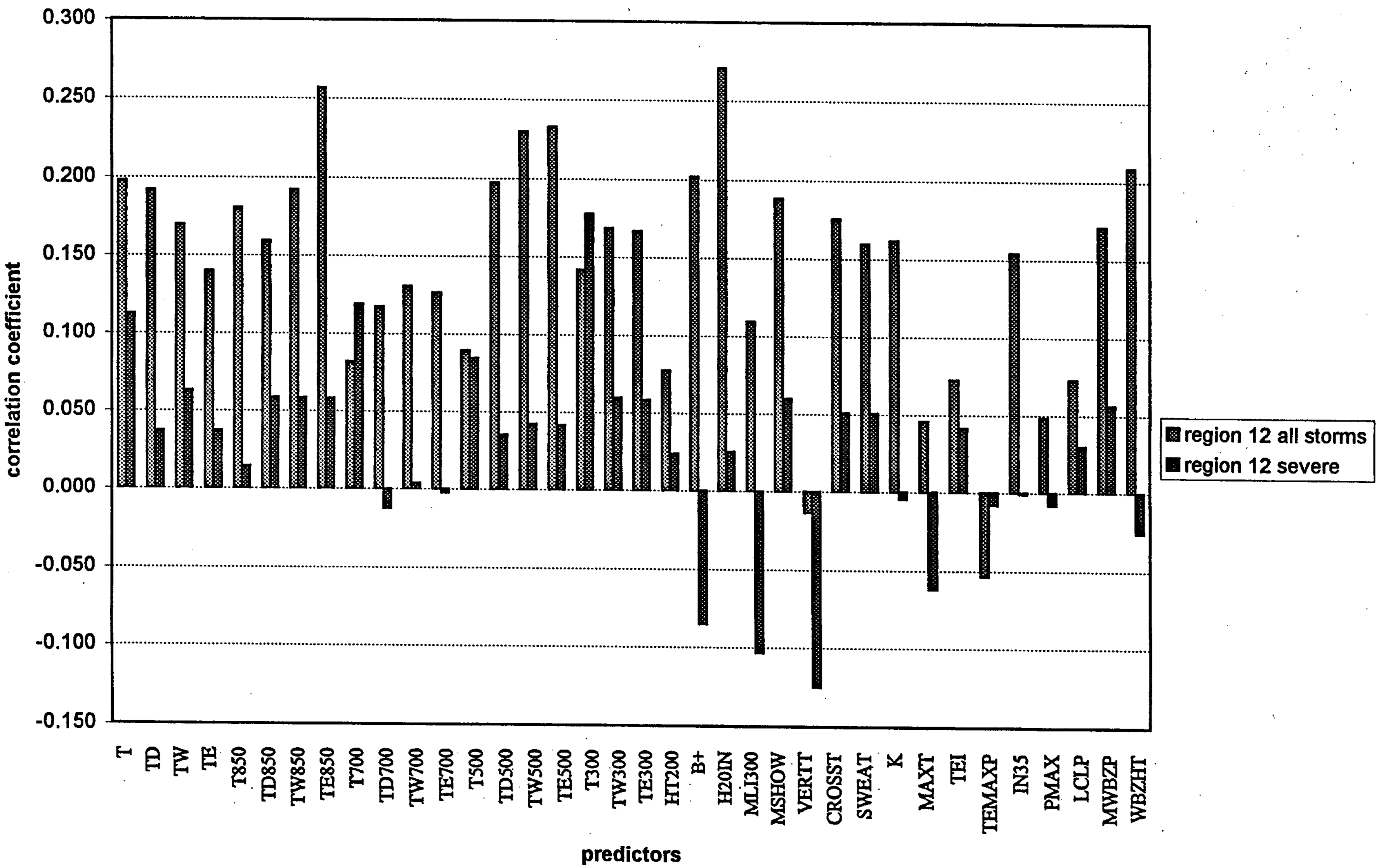


Figure 57. Same as Figure 55 but for 33 Statute Mile Radius.

Figure 33 is deceptive because there were no severe storm cases for the 1992 part of the data in this predictand set; in both 1992 and 1993, all-storm correlations are being compared with only 1993 severe storm data. Thus, it may not be unusual to see storm correlation coefficients higher than for the overall storm case. By the same token, there is the same lack of 1992 predictands for Figure 32; however, here the all-storm coefficients happen to be higher than the severe storm coefficients. Because of the lack of severe storm data, no significant conclusion should be drawn.

There is a similar lack of 1992 severe storm predictands (again, refer to Figures 58-63) for the remaining CONUS regions and, by referring to Figures 58 and 59, it is seen that both 1992 all-storm and severe predictands are lower in number across-the-board, for all regions.

One additional explanation for the anomalies in the correlation coefficient charts is that the CONUS regions were not defined by the same number of stations: Region 12 has one and Region 4 has four, whereas Region 5 has 14 stations.

Table 7 shows the top ten predictors, by region and predictand type, that were identified by analyzing correlation coefficients. At this point, it is reasonable to ask, without having any test data available (forecast fields from a mesoscale model at the required forecast intervals), what model(s) should be developed, given the fact that no testing will be done, and no comparisons can be made among the three different model types (those developed using the three predictand types). Given this situation, it was decided that a very limited amount of tests would be accomplished, to ensure that a generalized model, developed from 1992 and 1993 data for the all storm predictand type, produced reasonable outcomes. More exhaustive follow-on work can be accomplished relatively easy, given that the predictand/predictor databases have been constructed, edited, and manipulated for ease of use.

Because of the way in which these data spreadsheets have been constructed, models can be constructed using individual stations. Thus, if one wants to investigate an overseas location which is climatologically similar to one of our 71 stations, a storm forecast equation can be tailored for that overseas station or area.

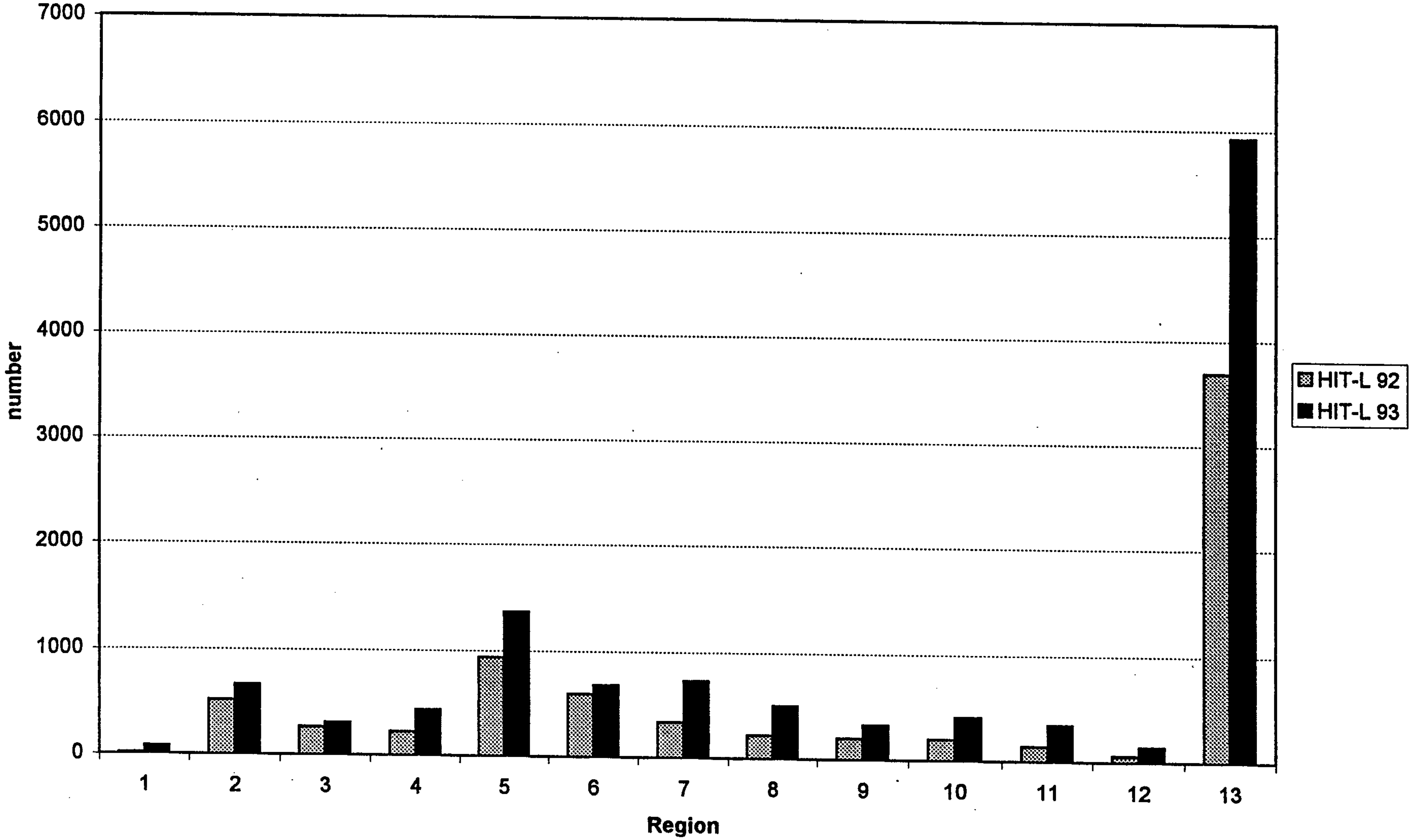


Figure 58. Comparison of Overall (100 Statute Mile) Predictands by Region and Year.

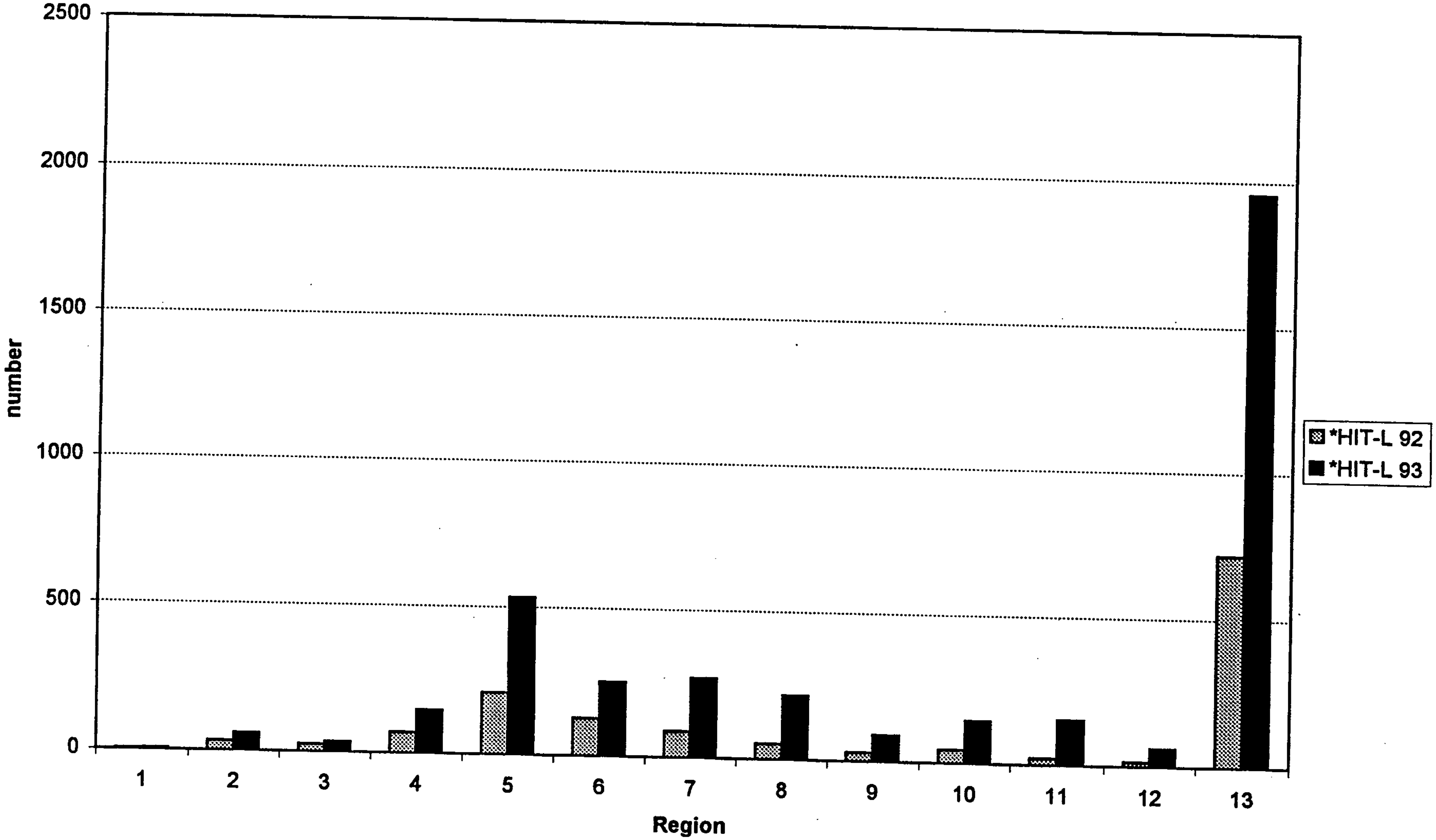


Figure 59. Same as in Figure 58 but for Severe Storm Predictands.

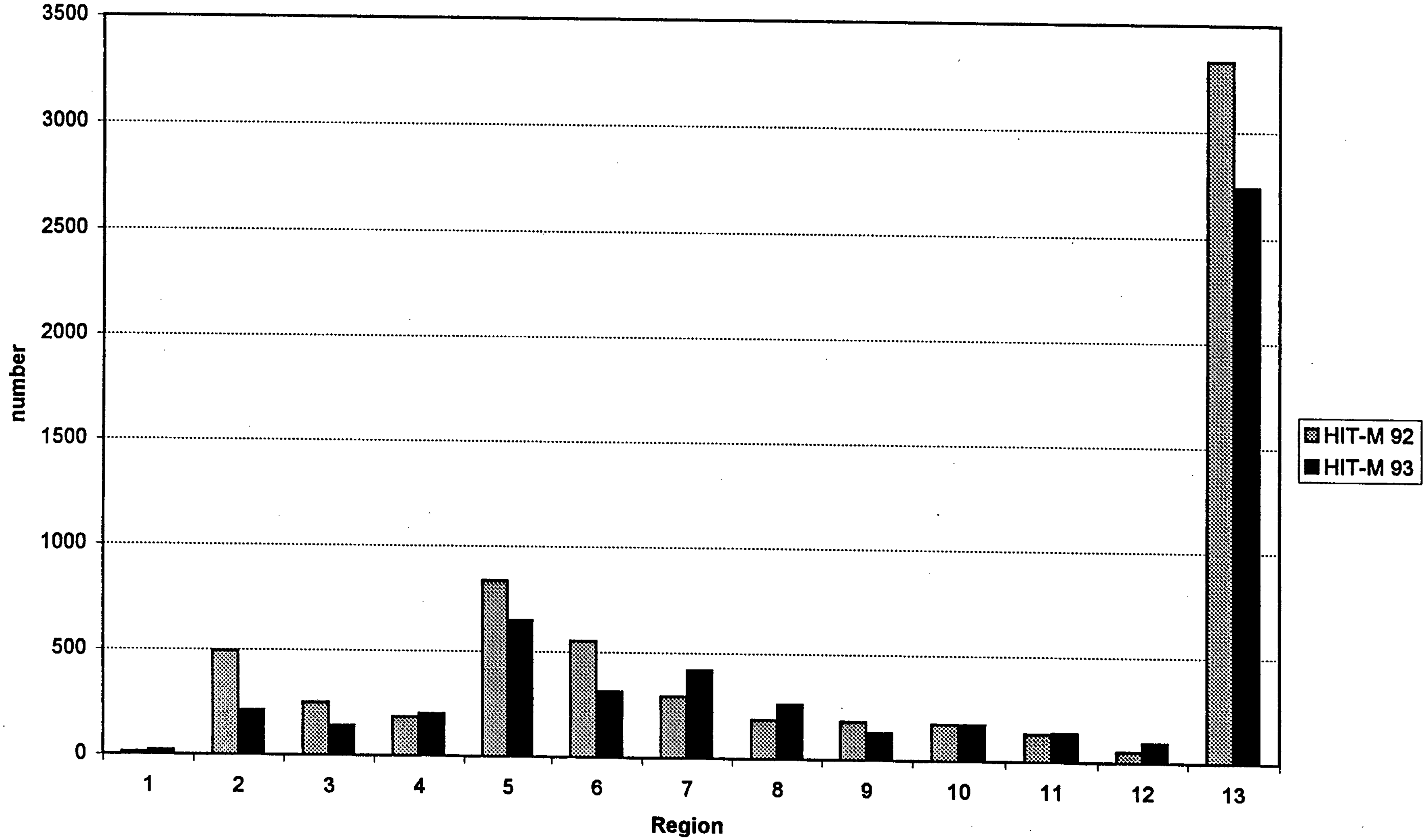


Figure 60. Same as in Figure 58 but for 66 Statute Mile Radius.

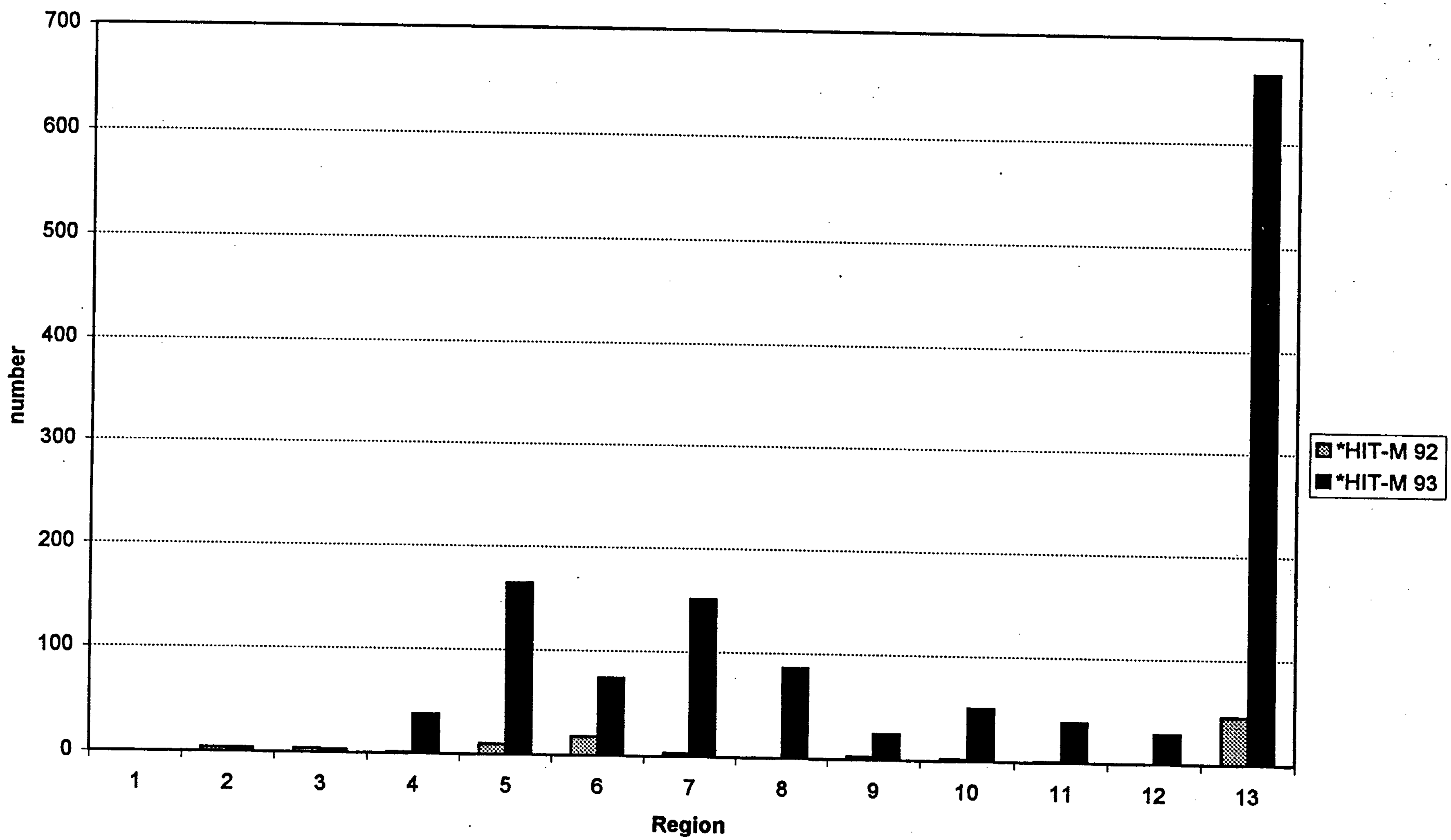


Figure 61. Same as in Figure 59 but for 66 Statute Mile Radius.

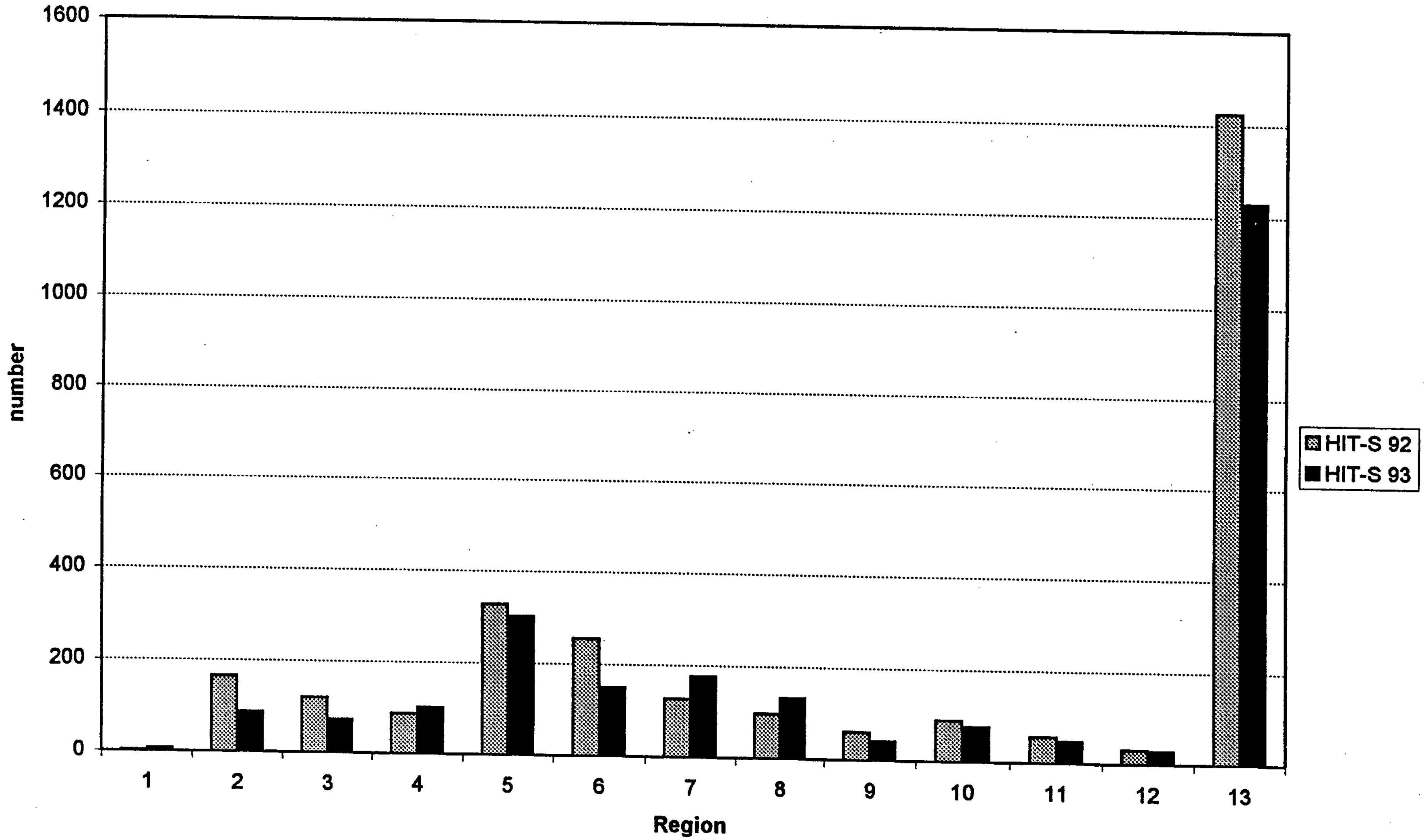


Figure 62. Same as in Figure 58 but for 33 Statute Mile Radius.

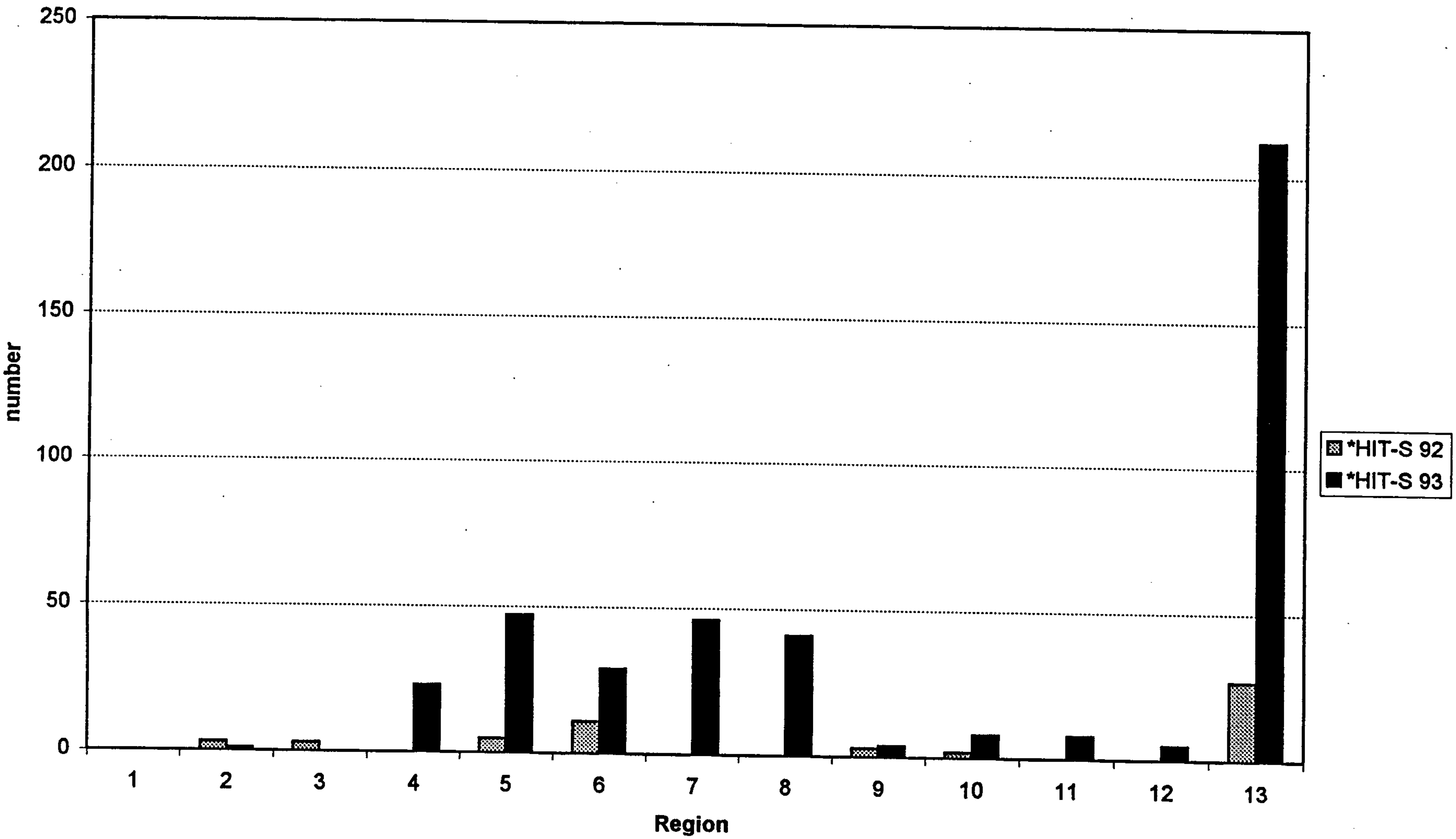


Figure 63. Same as in Figure 59 but for 33 Statute Mile Radius.

Table 7. Top Ten Predictors by Region and Predictand Type

REGION	PREDICTANDS	PREDICTORS									
		#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
1992-1993 ALL-33	1067	TE850	SHOW	K	TW850	TE700	TW	TW700	LI300	P-MAX	T850
1992-1993 SEV-33	212	TW700	TE850	LI300	TE500	TW	TW850	TW700	B+	TW500	TE
1992-1993 ALL-66	5197	TE700	TE850	TW850	SHOW	TW700	K	WBZ-P	TW	TW850	TD700
1992-1993 SEV-66	709	TE850	TE700	TW850	TW	TW700	TE500	LI300	WBZ-HT	WBZ-P	TW500
1992-1993 ALL-100	10907	K	SHOW	TW850	TE	TE850	TW	TD700	LCL-P	P-MAX	TD850
1992-1993 SEV-100	2596	TE850	TW850	TE	TW	TE700	SHOW	TD	TW700	LI300	TD850
REG 1 ALL-33	5	B+	LCL-P	P-MAX	LI300	K	TEI	T500	TE-MAXP	T	TW
REG 1 SEV-33	0										
REG 1 ALL-66	33	B+	K	TD700	T	SHOW	IN35	LCL-P	LI300	TW	TD500
REG 1 SEV-66	0										
REG 1 ALL-100	120	SHOW	K	IN35	CROSST	LCL-P	TE-MAXP	P-MAX	TD700	B+	TD850
REG 1 SEV-100	8	TE	K	LCL-P	SHOW	TD700	SWEAT	TW	CROSST	P-MAX	TD
REG 2 ALL-33	87	K	SHOW	TD700	P-MAX	LCL-P	TE	TD500	TW850	H20	TE850
REG 2 SEV-33	1	SWEAT	CROSST	IN35	TE-MAXP	K	VERT-T	TD700	SHOW	TD	TE
REG 2 ALL-66	631	K	SHOW	LCL-P	P-MAX	TD700	H20	TD850	TD500	TE	TE-MAXP
REG 2 SEV-66	8	TE-MAXP	CROSST	SWEAT	IN35	K	VERT T	TE	SHOW	TD	TD700
REG 2 ALL-100	1325	SHOW	K	TD700	LCL-P	P-MAX	TW850	TE850	TE	TD850	TD500
REG 2 SEV-100	85	SHOW	SWEAT	K	TE	TD850	P-MAX	LCL-P	H20	TE850	TD500
REG 3 ALL-33	69	T	TE700	K	TW700	TW	T850	SHOW	IN35	MAX T	TW850
REG 3 SEV-33	0										
REG 3 ALL-66	309	T	TE700	WBZ-HT	TW	TW700	MAX T	T850	IN35	WBZ-P	VERT T
REG 3 SEV-66	7	TE500	TW500	TD500	H20	TE300	TW300	T300	HT300	TW	TE
REG 3 ALL-100	665	K	SHOW	TW	IN35	H20	P-MAX	TE	LCL-P	TD700	TD
REG 3 SEV-100	54	K	TD700	SHOW	H20	TW	TD500	IN35	P-MAX	TE700	TW850
REG 4 ALL-33	82	K	TD700	TE700	T	TW700	WBZ-P	LI-300	B+	TW	IN35
REG 4 SEV-33	23	TD700	TE700	T	TW700	WBZ-P	LI-300	B+	TW	IN35	WBZ-HT
REG 4 ALL-66	341	K	TE700	TD700	TW700	WBZ-HT	H20	WBZ-P	TW500	TW	T
REG 4 SEV-66	39	H20	TD500	TD700	TW500	K	WBZ-HT	TW	TE500	TE700	TD
REG 4 ALL-100	823	K	SHOW	IN35	TD700	SWEAT	CROSST	P-MAX	WBZ-P	TW700	TE700
REG 4 SEV-100	213	K	SWEAT	LI-300	TE700	SHOW	TW700	IN35	TE850	TW	TE
REG 5 ALL-33	266	TE850	TW	TW850	TE700	SHOW	P-MAX	TE	SWEAT	TW700	LCL-P
REG 5 SEV-33	47	B+	LI-300	TE700	TE	TW	TE850	TEI	TE300	TD	TW700
REG 5 ALL-66	1320	WBZ-P	TW850	TE700	TE850	T850	WBZ-HT	TW700	TW	MAXT	TE
REG 5 SEV-66	175	LI-300	T850	TE850	TE700	B+	MAXT	TW	TW850	TW700	WBZ-HT
REG 5 ALL-100	2506	TW	TW850	TE850	TE700	WBZ-P	WBZ-HT	TD	TW700	SHOW	TE
REG 5 SEV-100	731	TE850	TW850	TE	TW	LI-300	TD	TE700	SWEAT	TEI	TW700
REG 6 ALL-33	136	TE	TW	TE850	LI-300	TW850	TE700	K	TD	SHOW	TW700
REG 6 SEV-33	29	TE500	TW500	TE850	TD	TE700	LI-300	WBZ-HT	T300	TW	HT200
REG 6 ALL-66	748	TE850	TW850	TE700	TD	TW700	WBZ-P	WBZ-HT	TW	TE	TW500
REG 6 SEV-66	92	TE850	TW850	LI-300	TD	TW	WBZ-HT	TE700	TE500	WBZ-P	T300
REG 6 ALL-100	1337	K	SHOW	TW850	IN35	TE850	TD850	TD700	TE	TW	TD
REG 6 SEV-100	361	LI-300	H20	TE850	TW	TE	TW850	TE700	TW700	WBZ-HT	TD
REG 7 ALL-33	151	H20	SHOW	K	TE850	TW850	IN35	CROSST	TD850	TD500	TE700
REG 7 SEV-33	46	LI-300	K	SHOW	H20	TD500	WBZ-P	TE500	TW500	IN35	MAX T
REG 7 ALL-66	516	H20	TE500	K	TW500	WBZ-P	TE700	TW700	WBZ-HT	TE850	TW850

Table 7. continued

REG 7 SEV-66	152	H20	WBZ-HT	WBZ-P	TE700	TE500	TW700	TW500	TE850	K	TE300
REG 7 ALL-100	1071	H20	K	TD700	TW850	TE850	TW700	TE700	WBZ-P	WBZ-HT	SHOW
REG 7 SEV-100	320	H20	TE850	TW850	TEMAX-P	TW700	TE700	K	WBZ-HT	WBZ-P	TD700
REG 8 ALL-33	119	H20	LI-300	TE700	TE850	TW700	K	WBZ-HT	TE850	B+	WBZ-P
REG 8 SEV-33	41	T	TE850	TW850	T850	WBX-HT	H20	TE500	TE850	T700	TW
REG 8 ALL-66	371	LI-300	TW	H20	WBZ-HT	TD	TE	T	TE700	TW700	TE850
REG 8 SEV-66	86	TE850	H20	T850	TW850	T	WBZ-HT	TE500	LI-300	TW500	TW
REG 8 ALL-100	959	H20	K	WBZ-P	WBZ-HT	TD700	TW850	TE850	TW700	TE700	SHOW
REG 8 SEV-100	271	TE850	MAXT	TW850	T850	H20	TW	LI-300	SHOW	TW700	K
REG 9 ALL-33	29	LCL-P	P-MAX	TE700	SWEAT	TE850	TE	MAXT	TW850	TD	WBZ-P
REG 9 SEV-33	4	LI-300	TE300	TE700	T300	TW300	T500	LCL-P	HT200	TE500	TD700
REG 9 ALL-66	280	TD	TW	TE850	TE700	WBZ-HT	TW850	TW700	SWEAT	WBZ-P	TD850
REG 9 SEV-66	28	TE300	TW300	T300	TE500	TE700	SWEAT	LCL-P	TE850	P-MAX	TW500
REG 9 ALL-100	571	SWEAT	TE850	SHOW	TW850	TW	TE	TD	K	TE700	WBZ-HT
REG 9 SEV-100	115	P-MAX	LCL-P	TE850	TE	TW	TD	TW850	H20	WBZ-HT	SWEAT
REG 10 ALL-33	60	LI-300	P-MAX	SWEAT	B+	T	LCL-P	TE850	TW	TW850	TE
REG 10 SEV-33	8	T	TW	T850	TE850	LI-300	TD	TE	TW850	SWEAT	TEI
REG 10 ALL-66	288	T	P-MAX	LI-300	TW	B+	TE850	LCL-P	TEI	TW850	TE700
REG 10 SEV-66	53	LI-300	T	B+	TEI	TW	P-MAX	TE850	T850	TW850	MAX T
REG 10 ALL-100	724	TE850	TW	TE	TW850	SHOW	P-MAX	LCL-P	TD	T	T850
REG 10 SEV-100	187	LI-300	TE850	TW	TE	TW850	T850	TEI	P-MAX	TD	MAX T
REG 11 ALL-33	37	LI-300	B+	SWEAT	SHOW	H20	TE850	P-MAX	TW850	K	IN35
REG 11 SEV-33	8	WBZ-HT	H20	WBZ-P	B+	SHOW	TE700	K	SWEAT	TW700	TE850
REG 11 ALL-66	237	T850	TE850	WBZ-HT	TW850	LI-300	WBZ-P	TE700	TW700	H20	TW
REG 11 SEV-66	40	TE700	TW700	LI-300	H20	TW	TD700	MAX T	TEI	TD	T
REG 11 ALL-100	567	H20	TE850	SHOW	TW850	P-MAX	TE	LCL-P	TW	TE	K
REG 11 SEV-100	174	TW	TE	TE850	LI-300	H20	T	TW850	TD	MAX T	T850
REG 12 ALL-33	26	H20	TE850	TW850	TE500	TW500	TW	WBZ-HT	TD	TE	TD850
REG 12 SEV-33	5	T300	T700	T	T500	TW	SHOW	TW300	TD850	TW850	TE850
REG 12 ALL-66	123	TE850	TW850	TD850	TW	IN35	WBZ-HT	CROSST	TD	T	T850
REG 12 SEV-66	29	T	TW	TD	TE	TE MAXP	T500	T850	TW850	TE850	T300
REG 12 ALL-100	239	TW	TE	TD	LI-300	MAX T	H20	T850	LCL-P	P-MAX	K
REG 12 SEV-100	77	T	TW	TE	TD	T850	MAX T	LI-300	TEI	T500	P-MAX

There are other possibilities on a larger scale–for example, a set of our CONUS stations could be constructed as a "region" of its own, apart from our 12 regions, to "match" any overseas region desired which exhibits similar frequency of occurrence of thunderstorm activity.

The next section recaps the limited testing accomplished to ensure the linear model we constructed produced reasonable probabilities for thunderstorm occurrence.

5. PREDICTION EQUATIONS

Tables 8 and 9 show the results of linear model construction for the 100 s.m. predictand and step regression, using S-PLUS advanced data analysis software. Only one predictor, PMAX, dropped out of the 1993 model as a single term deletion from the step regression. These models and those for the 66 s.m. and 33 s.m. predictands are listed below:

1993	**100 s.m.**	**66 s.m.**	**33 s.m.**
% Probability of Occurrence	= -0.00563*K +0.03872*MSHOW -0.04495*TW850 -0.00053*TE +0.00314*TE850 +0.02538*TW +0.01673*TD700 +0.00021*LCLP +0.00727*TD850 -0.15254	= -11.30133 +0.02185*TE700 +0.01442*TE850 -0.06043*TW850 +0.01431*MSHOW -0.03967*TW700 +0.00362*K +0.02166*TW	= -9.33211 +0.00625*TE850 +0.01078*MSHOW -0.02918*TW850 +0.02365*TE700 +0.00954*TW -0.04191*TW700 +0.00290*MLI300 -0.00411*T850

1992	**100 s.m.**	**66 s.m.**	**33 s.m.**
% Probability of Occurrence	= -0.00285*K +0.03662*MSHOW -0.04425*TW850 -0.00776*TE +0.00839*TE850 +0.01299*TW +0.01102*TD700 +0.00071*LCLP +0.00073*PMAX -4.06986	= -6.25947 +0.02258*TE850 -0.06384*TW850 +0.00146*K +0.00043*MWBZP +0.00487*TW	= -6.97971 +0.02057*TE850 +0.00627*MSHOW -0.00071*K -0.06408*TW850 +0.00292*TE700 -0.00496*TW +0.00233*MLI300

Table 8. 1993 Linear Storm Forecast Model Regression Results

Coefficients	Predictors Standard Value
(Intercept)	-0.1526
K	-0.0059
MSHOW	0.0391
TW850	-0.0459
TE	-0.0005
TE850	0.0032
TW	0.0252
TD700	0.0171
LCLP	0.0002
PMAX	0.0001
TD850	0.0078

Table 9. 1992 Linear Storm Forecast Model Regression Results

Coefficients	Predictors Standard Value
(Intercept)	-4.0699
K	-0.0029
MSHOW	0.0366
TW850	-0.0443
TE	0.0077
TE850	0.0084
TW	-0.0130
TD700	0.0110
LCLP	-0.0007
PMAX	0.0007
TD850	-0.0021

Table 10 summarizes a 10-sounding test of the 1993 and 1992 models. Data was taken from both 1992 + 1993 and so contains dependent and independent data. Figure 64 illustrates the similarity in the results, although the 1993 percentages are higher in almost every case, as might be expected. This is not the case for the other types of predictands. From this small test sample, it appears that the 100 s.m. predictand would provide a reasonable storm nowcast.*

A linear model was also created and limited tests conducted for maximum storm cloud top, using equilibrium level (EL) height as the predictand. The model is described below, and was built using 1993 data only for those soundings when EL height was computed by the SHARP program. Figure 65 shows the overall best predictors for EL height. MPLHT was dropped, because it produced ambiguous results for severe storm soundings when MPLHT was

Table 10. Summary of Limited Test for Thunderstorms

Test Sounding No.	1	2	3	4	5	6	7	8	9	10
K	20.00	9.00	3.00	42.00	36.00	-5.00	42.00	9.00	45.00	1.00
MSHOW	-4.00	0.00	-12.00	5.00	1.00	-10.00	3.00	-8.00	9.00	-13.00
TW850	11.70	9.10	3.90	17.90	13.70	8.70	17.50	5.60	18.90	3.20
TD	23.20	17.20	10.00	22.20	22.00	9.20	23.20	10.60	25.50	11.20
TE	354.47	334.01	316.87	358.33	348.46	328.59	354.11	313.64	376.14	315.79
TE850	330.60	322.80	308.80	353.50	337.40	321.40	351.90	313.00	357.90	307.10
TW	24.70	19.60	14.50	23.90	23.10	17.60	24.00	13.50	28.20	13.70
TD700	-4.60	-8.80	-17.00	5.70	4.00	-23.00	8.40	-9.80	5.40	-3.40
LCLP	945.45	243.51	245.88	892.34	958.98	176.60	828.90	279.54	854.09	175.18
TD850	7.80	6.60	3.30	17.60	13.40	1.40	16.00	-6.20	16.60	-15.40
PMAX	1017.00	300.00	300.00	974.00	1011.00	300.00	900.00	300.00	1000.00	300.00
MWBZP	-672.00	-749.00	-797.00	-609.00	-64.60	-745.00	-545.00	-729.00	-640.00	-780.00
TE700	325.77	322.19	308.47	337.68	333.14	314.40	345.35	316.87	342.27	324.84
TW700	1.90	0.50	-5.60	6.20	4.70	-2.80	8.60	9.00	7.70	1.60
MLI300	8.00	-1.00	-1.00	6.00	7.00	-2.00	9.00	0.00	18.00	-2.00
T850	17.80	12.60	4.60	18.60	14.20	17.40	21.00	15.80	24.60	14.60
STN	EYW	EYW	SLE	PIT	BRO	EDW	EDW	OAK	DDC	DDC
YEAR	1993	1993	1993	1993	1992	1992	1992	1992	1992	1992
MONTH	7	5	5	7	5	7	7	5	7	9
DAY	13	12	23	5	5	1	9	16	23	10
TIME(Z)	12	12	0	0	0	0	0	0	0	12
% Probability										
93 Model (100 s.m.)	71	32	15	94	90	4	90	24	112	21
92 Model (100 s.m.)	38	12	0	70	53	2	63	14	96	22
93 Model (66 s.m.)	35	15	3	59	48	1	64	31	80	8
92 Model (66 s.m.)	32	23	20	50	37	20	51	22	95	20
93 Model (33 s.m.)	16	12	0	31	25	0	37	0	41	3
92 Model (33 s.m.)	15	6	2	33	22	1	34	5	44	6
Note:	severe wx case	non-storm case	non-storm case	severe case	severe case	non-storm case	storm case	non-storm case	svr storm; LI = 15,	non-storm case
Overall Model (100 s.m.)	54	32	20	78	67	13	76	24	93	25
Overall Model (svr (100 s.m.))	18	8	1	36	24	0	40	1	48	2

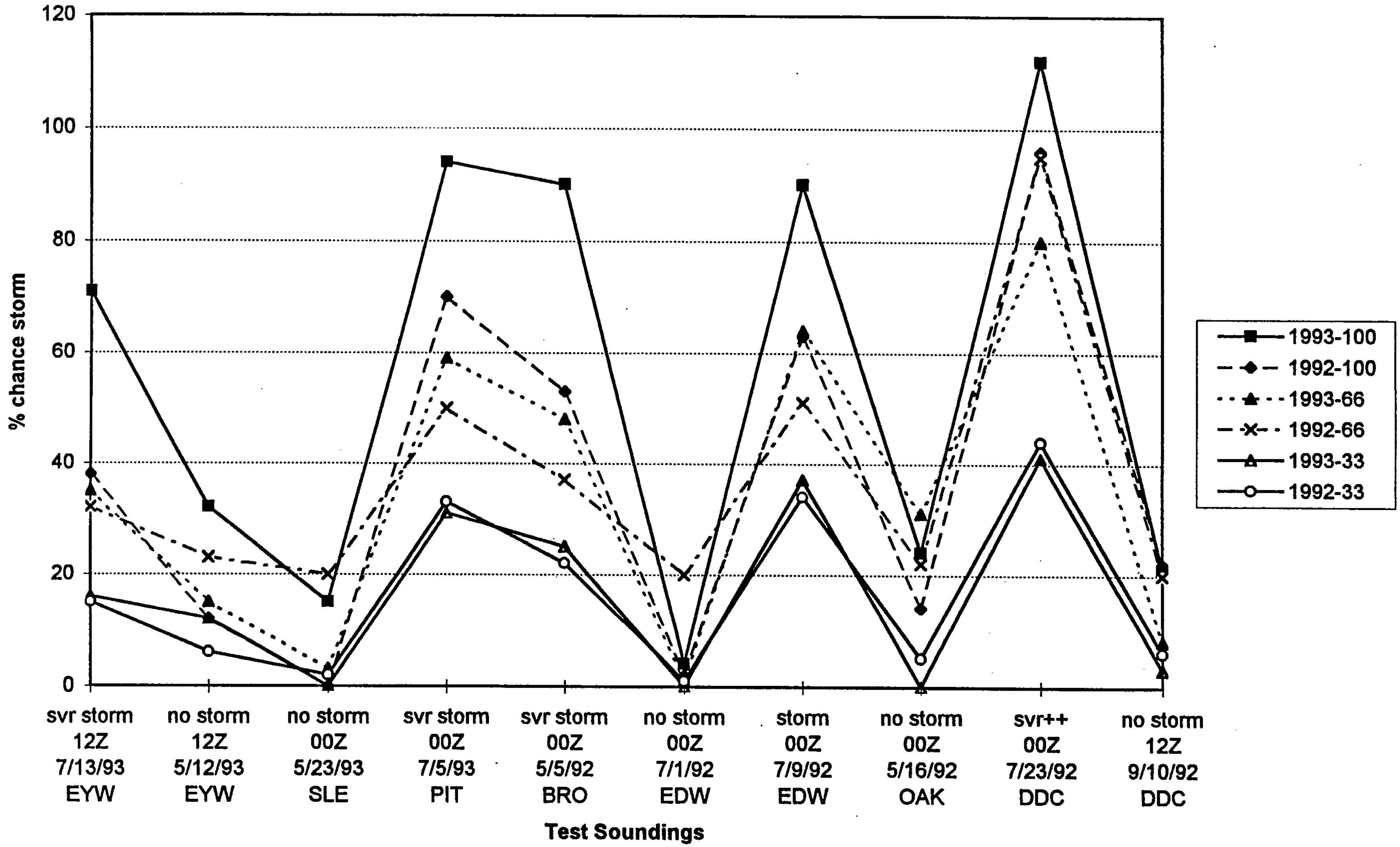


Figure 64. Limited Test Results.

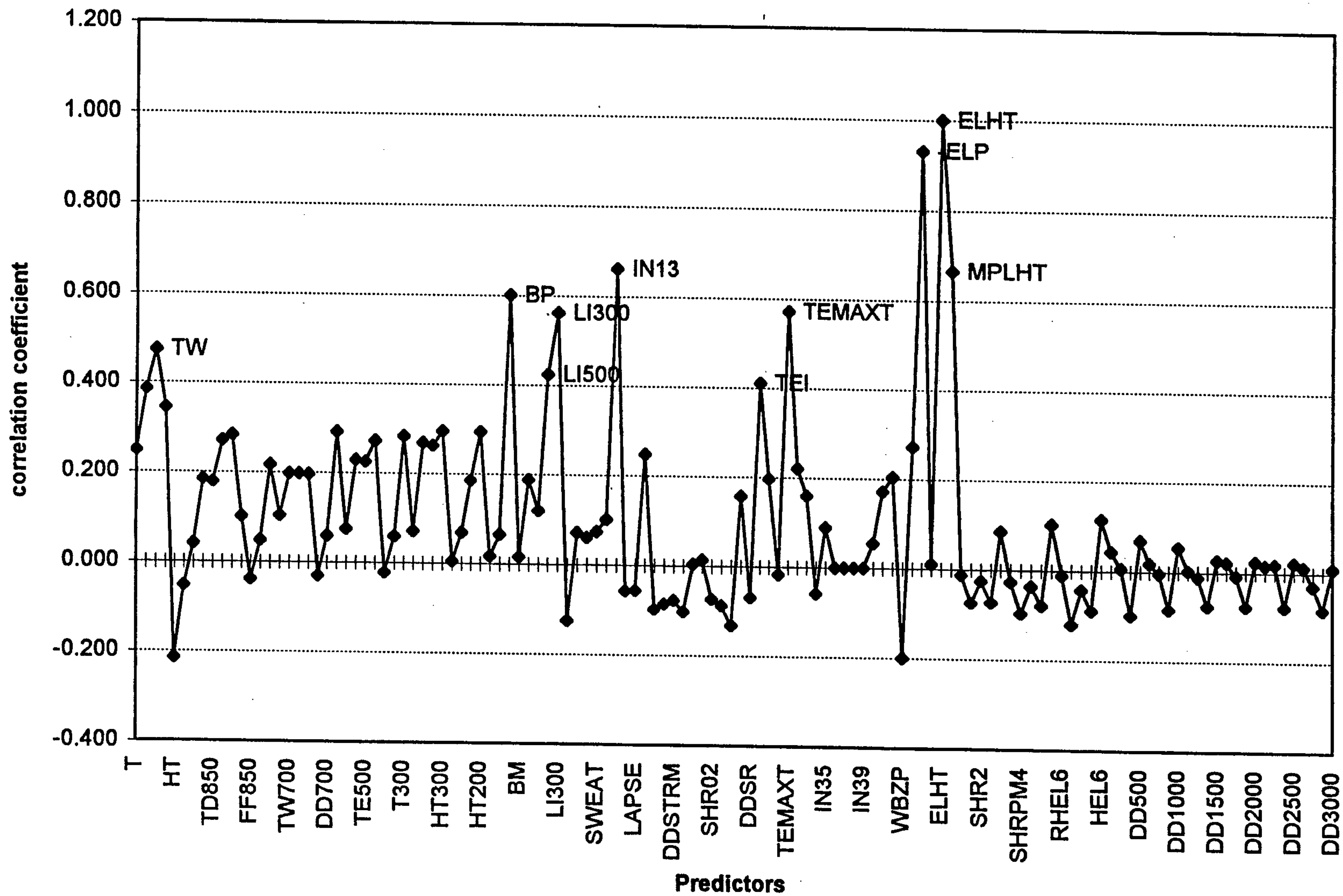


Figure 65. Correlation Coefficients for Maximum Storm Cloud Top Height Predictors.

between 0-10,000 feet (approx. limits); and IN13 was dropped due to our inability to define this parameter.

height =	0.8084105	BP
	-493.7144	LI300
	507.6571	TW
	-39.84367	TEI
	148,9849	TEMAXT
	-25504.33	

The results in Table 11 are promising, since the model produced 30,000+ foot cloud tops for the severe storm cases. Test sounding #5 and #7 are anomalous; however, they do not necessarily invalidate the model, since (in the case of #7) storm potential is not always realized and (in the case of #5) thunderstorm tops of 20-25,000 feet are not uncommon. Linear model summary and step regression details are shown in Table 12.

Table 11. Summary of Limited Test for Maximum Cloud Top Height

Test Sounding #	1	2	3	4	5	6	7	8	9	10
BP	472	424	1907	0	236	1629	1119	15	6054	33
LI300	1	-2	-7	0	0	-7	-3	1	-18	0
TW	20.3	15.6	23.1	5.8	17.2	17.3	22.4	14.4	28.2	14.6
TEI	8.87	25.88	18.6	19.73	14.95	23.7	17	15.21	50.76	6.45
TEMAXT	338.8	339.64	350.36	333.13	327.2	351.08	352.86	329.59	376.14	325.6
STN	VBG	VBG	DDC	DDC	OMA	OMA	JAN	LZK	DDC	RAP
YR	1992	1992	1992	1992	1992	1992	1992	1992	1992	1992
MO	7	9	9	5	5	7	7	5	7	9
DAY	13	14	7	6	15	2	26	3	23	13
TIME(Z)	0	0	12	12	00	12	12	12	00	12
Height	24,506	25,396	30,951	23,340	22,838	30,630	28,775	22,511	42.293	22,775
Note	non-storm case	non-storm case	svr-storm case	non-storm wx case	storm case	svr-storm case	Non-storm case	non-storm case	svr storm; L = -15, SWEAT = 595	non-storm case

Table 12. 1993 Linear Storm Cloud Top Model Regression Results

Coefficients	Predictors Standard Value
(Intercept)	-25504.3300
BP	0.8084
LI300	-493.7144
TW	507.6571
TEI	-39.8437
TEMAXT	148.9849
STD ERROR	3955.5300

As stated in the Introduction, prediction equations would be used to satisfy Air Force Weather's Theatre Battle Management stated requirements and these requirements included the development of present weather prediction techniques, an effort running in conjunction with the thunderstorm prediction effort described here. Researchers working on present weather prediction techniques identified a need for a generalized thunderstorm and severe thunderstorm equation based on 100 statute mile radius but for ±2 hours in time. To satisfy this need, generalized thunderstorm equations were developed using the 1992 and 1993 data bases and scheme described previously but limiting severe thunderstorm occurrence to ±2 hours as derived from the Storm Summary data base and including the "Lightning –s" predictands (see Figure 4). The resulting General Thunderstorm Global Algorithm and Global Severe Thunderstorm Equation, which were incorporated in the present weather prediction effort, is shown in Tables 13 and 14.

Table 13. General Global Thunderstorm Linear Storm Forecast Model Regression Results

COEFFICIENTS	PREDICTORS STANDARD VALUE
(INTERCEPT)	+127.3597
LATITUDE	+0.01219
LONGITUDE	-0.03100
T	+0.10387
TD	+0.05580
TW	-0.18570
TE	+0.06289
T850	-0.05212
TD850	-0.02513
TW850	+0.67128
TE850	-0.10506
T700	-0.14569
TD700	-0.05913
TW700	+0.61291
TE700	-0.10503
HT700	+0.00135
T500	-0.48531
TD500	-0.02522
TW500	+0.23918
TE500	+0.00323
HT500	-0.00575
T300	-0.01009
TD300	+0.25248
TW300	+0.24882
TE300	-0.13560
HT300	-0.00025
HT200	+0.00035
BPLUS	-0.00010
H20	+0.68911
VERTT	+0.55158
CROSST	+0.45551
SWEAT	-0.00008
IN13	+0.00071
CAP	-0.01504
MAXT	-0.04546
dd06	+0.00080
DDSTRM	-0.00262
TEI	-0.00358
TEIPB	-0.00008
TEMAXT	+0.01839
TEMAXP	+0.00021
TEIMINT	-0.01869
IN35	-0.53530
PMAX	+0.00088
LCLP	-0.00081
WBZP	-0.00130
WBZHT	+0.00001
ELP	-0.00086
MPLP	+0.00016
ELHT	+0.000004

MPLHT	+0.000005
SHR4	-0.05899
HEL4	-0.00095
SHR6	+0.02224
HEL6	-0.00101
F1000	+0.04607
FF1500	-0.07350
DD1500	-0.00046
FF2000	+0.10574
DD2000	+0.00008
FF2500	-0.10079
DD2500	+0.00024
FF3000	+0.00052
DD3000	+0.00062

Table 14. Severe Global Thunderstorm Linear Storm Forecast Model Regression Results

COEFFICIENTS	PREDICTORS STANDARD VALUE
(INTERCEPT)	+14.889738
TW	+0.112705
TD850	-0.353522
TW850	+0.730811
TE850	-0.090695
BPJDKG	+0.000105
CROSST	+0.307847
SWEAT	-0.000272
FFOT6	+0.003989
FFSTRM	+0.038964
SHRBRN	+0.014673
IN35	-0.055614
SHR4	+0.056171
SHR6	-0.000667

*DETERMINANT THRESHOLD: > .08 = SEVERE THUNDERSTORM: < .08 = NO SEVERE THUNDERSTORM. POD = 0.627. FAR = 0.230.

6. PRELIMINARY TEST USING RAOB DATA

Figure 66 test results appear promising. To draw a more confident conclusion from these limited results more information about the test soundings and the associated local and synoptic weather patterns would be helpful. Below is some background information:

Storm Cases

Svr storm, 12Z 7/13/93, EYW: At least two funnel clouds were sighted, and 1.5" diameter hail was reported over South Florida locations this day. KI = 20; Showalter Index = 5; Lifted Index = -8.

Svr storm, 00Z 7/5/93, PIT: Severe thunderstorm winds were reported. KI = 42; Showalter Index = -5; Lifted Index = -6.

Svr storm, 00Z 5/5/92, BRO: 3/4" diameter hail and wind gusts to 60 mph were reported. At 1256CST, a waterspout was reported in Galveston City. KI = 36; Showalter Index = -1; Lifted Index = -7.

Storm, 00Z 7/9/92, EDW: Tropical moisture associated with Tropical Storm Darby moved into southern California and brought unseasonal rains to the area, ranging from a trace to 3/4" in the mountains. KI = 42; Showalter Index = -3; Lifted Index = -9.

Severe storm, 00Z 7/23/92, DDC: Severe storms dropped large hail (up to 1.75" diameter in surrounding areas) and heavy flooding rainfall. KI = 45; Showalter Index = -9; Lifted Index = -18.

No storm, 12Z 5/12/93, EYW: No storm activity reported from 8-18 May 93. KI = 9; Showalter Index = 9; Lifted Index = 1.

No storm, 00Z 5/23/93, SLE: No storm activity from 20-24 May 93. KI = 3; Showalter Index = 12; Lifted Index = 1.

No storm, 00Z 7/1/92, EDW: No storm activity in the area for entire month of June and first half of July 92. KI = 5; Showalter Index = 10; Lifted Index = 2.

No storm, 00Z 5/16/92, OAK No storm activity in the area for entire month. KI = 9; Showalter Index = 8; Lifted Index = 0.

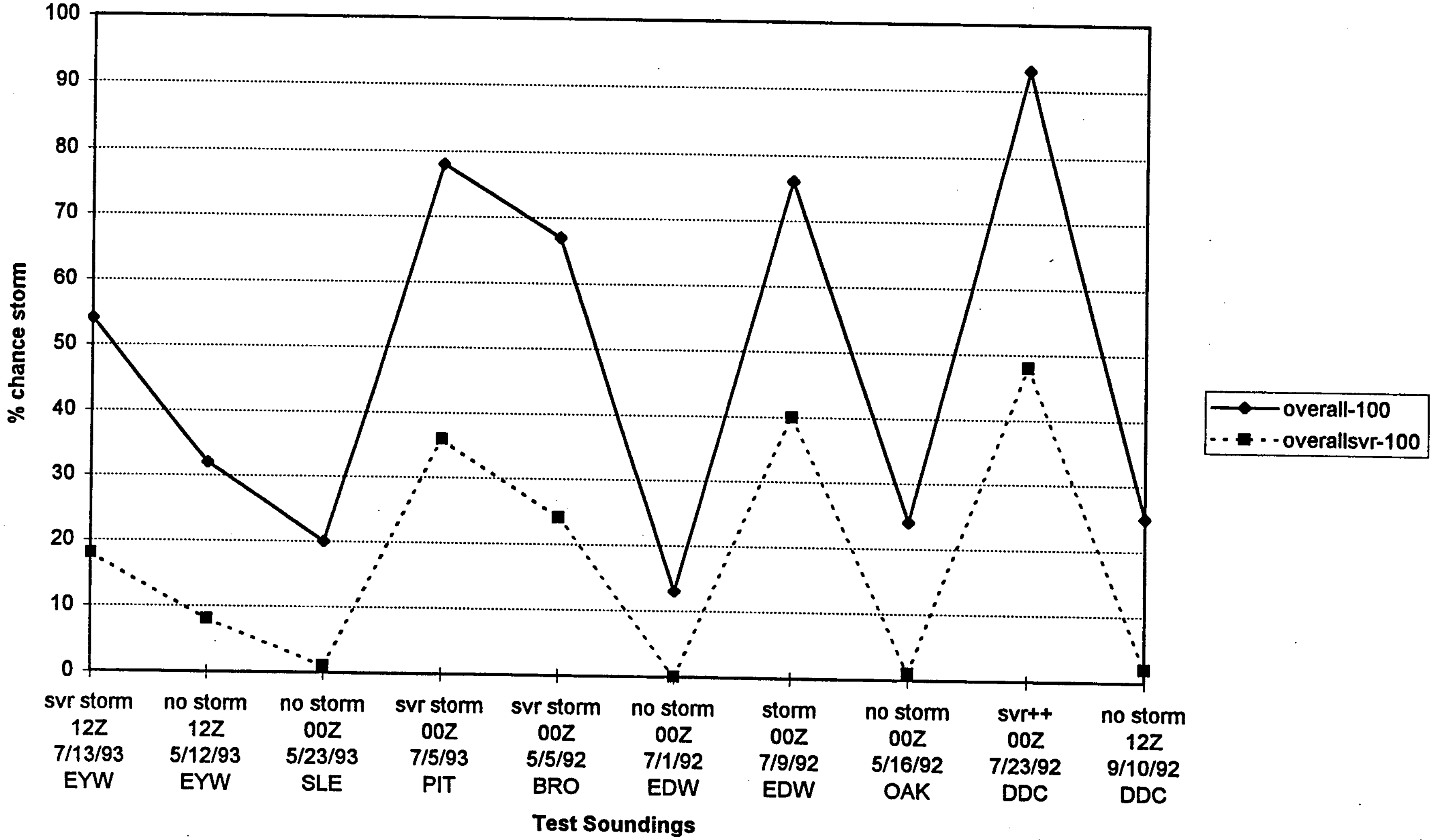
No storm, 12Z 9/10/92, DDC: No storms in the area for ~ 10 days. KI = 1; Showalter Index = 13; Lifted Index = 2.

In retrospect, more testing would help substantiate these limited but promising results. Additionally, forecast predictors would help to tailor this study's models for forecast periods of interest.

Now that we have an idea of the general character of 1992 and 1993 thunderstorms over our CONUS regional network with respect to our chosen set of predictors, and have accomplished some limited tests, we did the following: 1) developed generalized (1992-1992) models to diagnose ("nowcast") storm and severe storm occurrence (see Table 13); 2) conducted limited testing of both models (see Figure 66); 3) developed similar "generalized" models for each of the 12 CONUS regions (see Table 14).

Table 15. Generalized (1992-1993) Overall and Severe Storm Models for all Regions at 100 Statute Mile Radius

Coefficients	Region All/Type All	Region All/Type Severe
(Intercept)	-5.0264	-10.4451
K	-0.0047	
MLI300		0.0020
MSHOW	0.0235	0.0032
PMAX	0.0001	
TD		0.0053
TD700	0.0120	
TD850	0.0054	0.0014
TE	-0.0009	-0.0016
TE700		0.0128
TE850	0.0194	0.0226
TW	0.0193	0.0078
TW700		-0.0252
TW850	-0.0750	-0.0643
STD Error	0.4472	0.2959



66. Limited Test Results for Generalized Case (1992 and 1993).

Table 16. Generalized (1992 and 1993) Overall and Severe Storm Model Statistics for 12 CONUS Regions at 100 Statute Mile Radius.

PREDICTORS	#1/ALL	#2/ALL	#2/SEVERE	#3/ALL	#3/SEVERE	#4/ALL	#4/SEVERE	#5/ALL	#5/SEVERE
INTERCEPT	-0.42576	2.03584	0.46383	-2.64962	-1.67035	-17.94590	-22.42155	-8.14148	-15.13035
CROSST	0.01768								
H2OIN				0.41941					
K	0.01053	0.00255			0.01166	0.01077			
LCLP		0.00007		-0.00107					
MAXT									
MLI300									0.01075
MSHOW		0.02964	0.00402	0.01937	-0.01299		0.01908	0.01762	
MWBZP								0.00244	
PMAX			0.00002	0.00093		0.00016			
SWEAT			0.00018			0.00039	0.00047		0.00016
T500									
T850									
TD				-0.00938					0.00622
TD500		0.00335	0.00062		0.00135				
TD700					-0.01448				
TD850	-0.02622								
TE		0.00551	0.00125	0.01039			0.01171	-0.00108	-0.00142
TEMAXP	0.00026								
TE700					0.00423	0.05585	0.07521	0.02101	0.03821
TE850		-0.01046	-0.00262				-0.01655	0.01432	0.01062
TW			-0.04370				-0.00787	0.02634	0.00910
TW700						-0.14830	-0.18020	-0.03258	-0.08329
TW850								-0.06571	-0.02846
WBZHT								-0.00005	
STND ERROR	0.36307	0.43450	0.15260	0.44143	0.19256	0.46517	0.33364	0.45550	0.31878

Table 16. (continued)

PREDICTORS	#6/ALL	#6/SEVERE	#7/ALL	#7/SEVERE	#8/ALL	#8/SEVERE	#9/ALL	#9/SEVERE
INTERCEPT	-12.26650	-12.57954	-9.92838	-19.48234	5.50042	-27.71680	-5.07733	-8.08099
CROSST								
H2OIN		0.21165	0.45485	0.15358	0.19209			0.04478
K					0.02292	0.00786	0.00400	
LCLP								0.00033
MAXT						0.05235		
MLI300		0.00453						
MSHOW	0.02130							
MWBZP				-0.01062				
PMAX								-0.00239
SWEAT							0.00066	
T500								
T850						-0.05249		
TD		-0.01259					0.00756	
TD500								
TD700	0.00755				-0.02458			
TD850								
TE	0.00791	0.02093					-0.00093	0.00120
TEMAXP								
TE700			0.03016		-0.06952			
TE850	0.03526	0.02259		0.03073	0.05407	0.08164	0.01756	0.02542
TW		-0.05338					0.01963	0.00663
TW700			-0.07419		0.17757	-0.02543		
TW850	-0.12602	-0.07296		-0.08781	-0.16160	-0.25420	-0.05584	-0.06548
WBZHT				0.00027	0.00003			
STND ERROR	0.45689	0.33494	0.45220	0.35731	0.45807	0.35478	0.42264	0.22275

Table 16. (continued)

PREDICTORS	#10/ALL	#10/SEVERE	#11/ALL	#11/SEVERE	#12/ALL	#12/SEVERE
INTERCEPT	-13.70281	-11.64537	-10.47922	-14.08532	-1.42987	-0.80596
CROSST						
H2OIN					0.32159	
K						
LCLP						
MAXT		-0.02042		0.00427	0.01345	
MLI300		0.02198				
MSHOW	0.01727		0.01066			
MWBZP						
PMAX		-0.00011			0.00031	0.00024
SWEAT						
T500						0.02735
T850	0.01540	0.05287				
TD				-0.00644		-0.09803
TD500						
TD700						
TD850						
TE	-0.00191	-0.00311	0.03914	0.02125		
TEMAXP						
TE700						
TE850	0.04798	0.04525		0.02633		
TW	0.02453	0.01587	-0.10982	-0.04961		0.13467
TW700						
TW850	-0.14814	-0.12540		-0.06896		
WBZHT						
STND ERROR	0.43127	0.27036	0.44909	0.30258	0.44320	0.40375

7. ALGORITHM TEST USING MM5 FORECAST

A limited study using MM5 forecasts was performed on the thunderstorm algorithm models described earlier in this report. The study consisted of four main phases; generating forecasts with the Penn State/NCAR Mesoscale Model (MM5), using the output from MM5 for input to Skew-T, Hodograph Analysis and Research Program (SHARP) to produce the thunderstorm predictor values; using the spreadsheet generated by SHARP to compute thunderstorm probabilities with the regional and general thunderstorm models; and evaluate the forecast probabilities generated by the models with observed data.

7.1 MM5 Forecast

The MM5 system was used in its basic form. The only deviation from a normal MM5 run was to use MRF forecast files for the initial and boundary condition. The model produced the analysis, 12, 24, 36, and 48 hour forecasts for both the 0 and 12 Z initial times for the 31 days of August 1993. The whole suite of MM5 pre- and post- processors were used for this study: TERRAIN, DATAGRID, RAWINS, and INTERP (front-end and back-end). The inner domain was defined with a 46 by 61 grid with a resolution of 80 kilometers over the continental U.S. (CONUS). The hydrostatic version of the model was used in this study. The standard output files from the post-processor INTERP, executed in back-end mode, were used in the processing steps. This process generated 62 (2 * 31 days) output files. Most output files contained five forecast durations. There were a total of 304 forecasts out of a possible 310 forecasts (62 * 5 forecasts) when all the production runs were complete. There were six missing forecast durations from three model runs due to unknown model errors and there was not enough time to find and rectify them.

7.2 SHARP

Before running SHARP, the MM5 output had to be interpolated from its gridded format to the 69 RAOB site locations (Table 2) and converted into files that SHARP can read (Figure 1). A FORTRAN program was developed to interpolate to the RAOB sites and create files that mimic a RAOB record file. This process created 20976 files (304 forecasts * 69 RAOB stations).

A Microsoft Excel macro was then used to create the input files for SHARP with the newly created RAOB files. SHARP was then used to compute the predictors (Table 1) for the thunderstorm models. SHARP created five formatted files, one for each forecast duration, containing 4278 (62 * 69) records. These data files were created in column form for easy ingest to a spreadsheet.

7.3 Thunderstorm Probabilities

The data files created by SHARP were imported into Excel spreadsheets. The "all thunderstorms" models, both regional and general, and the "severe thunderstorms" models, both regional and general, were used to generate probabilities using the predictors supplied by SHARP.

At this point in the study, some concerns came to light. There are seven predictors that rely on quantities from the 850mb level. Some of the models use the temperature, dewpoint temperature, wetbulb temperature, and theta-E at the 850mb level. Also, some models use the K Index, SWEAT Index, and the Showalter Index values that are derived from 850mb quantities. Some RAOB station elevations were higher than the 850mb level. Regions two and three were comprised of stations that were higher. So the resulting probabilities were not valid due to missing predictors. These two regions were not evaluated any further. Also, the RAOB station for Denver was excluded from the evaluation of the region five models. Further study may be needed to consider elevation and the predictors that are needed.

7.4 Results of Forecast Experiments

We validated the forecasts of the probability of all thunderstorms and severe thunderstorms separately for seven of the regions. This was accomplished by comparing the diagnosed probability of thunderstorm occurrence at each station within a given region with the corresponding observation of thunderstorm occurrence at the forecast valid time. The validation measure used was the mean-squared accuracy of the diagnosed probability:

$$\overline{P} = \frac{1}{M}\sum_{m=1}^{M}(\widetilde{P}_m - O_m)^2$$

where:

$\widetilde{P}$ = Probability computed from the model

O = Storm (1)/ No Storm (0)

M= Number of observations (all stations in a region for all forecasts).

The smaller the value of $\overline{\mathrm{P}}$, the better the accuracy of the predicted probability. Table 17 lists the results of the validation of the predictions of the "all thunderstorms" diagnosis models against all thunderstorms, while the results given in Table 18 are a product of the validation of the "severe thunderstorms" models against just the severe thunderstorm occurrences (that is, $O_m = 0$ for non-severe thunderstorms). The column key for the tables is as follows: column A - probability computed with regional models; column B - probability computed with region 5 model using region 6 data / region 6 model using region 5 data; column C - probability computed with generalized models; column D - probability is zero; column E –probability is randomly generated; column F –probability randomly selected from probabilities generated by the regional models; column G –probability randomly selected from probabilities generated by the generalized models;

Column H –probability bias (remove the squared exponent from the above equation) based on probability computed with regional models.

In column A of the tables, we list the validation scores where the probability was computed from the diagnosis model developed for the respective region. The most notable trend in these scores, are the better skill shown for the severe thunderstorms in all regions. We attribute this to the fact that, in the data base, there were many fewer occurrences of severe thunderstorms than non-severe thunderstorms. Thus, a zero probability of severe thunderstorms is likely to be a good forecast, and it is much easier to obtain good forecasts without a skillful technique. This is shown to be true by an examination of column D, in which a zero probability was used in place of the computed probability in the validation. It is clear that, with the exception of region 8, the skill of the zero probability forecast is as good or better for severe thunderstorms than the model's skill. This is clearly not true for the all thunderstorms verification, in which there are many more occurrences of thunderstorms making a perpetual zero probability a bad forecast. We do not see a significant trend of skill change with forecast duration in either table. While there are some regions in which the greater forecast durations show worse scores, we doubt that the differences between these scores and those at smaller durations are significant. Some regions show better scores than others: for all thunderstorms, region 4 has the best scores while regions 6 and 11 are virtually tied for the worst. In contrast, for severe thunderstorms, region 11 has the best results while regions 4 and 6 are the worst.

Table 17. Evaluation of the All Thunderstorms Models

Forecast Hour	Region	A	B	C	D	E	F	G	H
0	4	0.190		0.208	0.620	0.327	0.233	0.255	-0.023
12	4	0.174		0.202	0.617	0.365	0.254	0.264	0.006
24	4	0.171		0.198	0.615	0.332	0.245	0.265	0.014
36	4	0.193		0.206	0.609	0.343	0.248	0.260	-0.016
48	4	0.192		0.198	0.601	0.312	0.276	0.257	-0.028
0	5	0.185	0.188	0.197	0.641	0.330	0.252	0.283	-0.017
12	5	0.188	0.188	0.190	0.636	0.338	0.259	0.261	0.014
24	5	0.187	0.189	0.193	0.643	0.347	0.268	0.271	-0.001
36	5	0.194	0.194	0.198	0.637	0.314	0.259	0.271	0.000
48	5	0.203	0.205	0.209	0.635	0.339	0.259	0.266	-0.008
0	6	0.212	0.213	0.211	0.650	0.342	0.241	0.238	0.054
12	6	0.226	0.224	0.213	0.650	0.312	0.245	0.236	0.105
24	6	0.224	0.216	0.214	0.661	0.313	0.244	0.241	0.086
36	6	0.228	0.226	0.222	0.650	0.345	0.241	0.243	0.054
48	6	0.225	0.228	0.222	0.650	0.357	0.245	0.231	0.019
0	7	0.179		0.162	0.825	0.336	0.228	0.185	-0.213
12	7	0.182		0.157	0.826	0.327	0.200	0.185	-0.217
24	7	0.188		0.161	0.835	0.305	0.233	0.189	-0.251
36	7	0.203		0.170	0.840	0.308	0.230	0.189	-0.277
48	7	0.205		0.170	0.852	0.301	0.230	0.189	-0.292
0	8	0.178		0.172	0.743	0.380	0.251	0.231	-0.193
12	8	0.154		0.161	0.740	0.349	0.281	0.237	-0.156
24	8	0.180		0.175	0.749	0.342	0.245	0.219	-0.187
36	8	0.203		0.182	0.748	0.326	0.273	0.212	-0.215
48	8	0.210		0.177	0.772	0.336	0.303	0.240	-0.272
0	10	0.200		0.186	0.625	0.330	0.278	0.269	-0.138
12	10	0.195		0.184	0.633	0.347	0.275	0.260	-0.137
24	10	0.205		0.194	0.653	0.342	0.286	0.270	-0.167
36	10	0.220		0.202	0.637	0.295	0.298	0.283	-0.181
48	10	0.228		0.208	0.650	0.336	0.284	0.275	-0.213
0	11	0.221		0.211	0.517	0.333	0.265	0.266	-0.014
12	11	0.221		0.209	0.522	0.339	0.260	0.259	-0.001
24	11	0.227		0.218	0.540	0.330	0.263	0.270	-0.021
36	11	0.229		0.224	0.514	0.316	0.272	0.292	-0.008
48	11	0.232		0.234	0.528	0.313	0.284	0.295	-0.033

Table 18. Evaluation of the Severe Thunderstorms Models

Forecast Hour	Region	A	B	C	D	E	F	G	H
0	4	0.135		0.120	0.140	0.349	0.163	0.134	0.143
12	4	0.163		0.123	0.140	0.317	0.181	0.135	0.197
24	4	0.164		0.130	0.146	0.318	0.189	0.140	0.182
36	4	0.153		0.123	0.135	0.317	0.168	0.135	0.152
48	4	0.142		0.115	0.128	0.326	0.160	0.123	0.137
0	5	0.115	0.116	0.082	0.083	0.326	0.133	0.092	0.176
12	5	0.129	0.134	0.084	0.082	0.325	0.147	0.098	0.199
24	5	0.127	0.130	0.085	0.083	0.311	0.149	0.098	0.194
36	5	0.123	0.124	0.081	0.077	0.338	0.138	0.092	0.196
48	5	0.121	0.123	0.084	0.082	0.322	0.137	0.097	0.181
0	6	0.142	0.160	0.125	0.139	0.332	0.160	0.131	0.174
12	6	0.166	0.179	0.130	0.144	0.312	0.175	0.142	0.209
24	6	0.164	0.173	0.134	0.149	0.322	0.176	0.146	0.195
36	6	0.151	0.161	0.125	0.138	0.353	0.156	0.131	0.175
48	6	0.132	0.145	0.113	0.122	0.335	0.145	0.122	0.161
0	7	0.108		0.108	0.107	0.337	0.123	0.114	0.127
12	7	0.104		0.108	0.106	0.343	0.124	0.112	0.108
24	7	0.100		0.104	0.110	0.351	0.113	0.116	0.088
36	7	0.095		0.098	0.099	0.348	0.107	0.106	0.085
48	7	0.101		0.103	0.105	0.352	0.102	0.108	0.077
0	8	0.107		0.111	0.134	0.300	0.134	0.135	0.054
12	8	0.107		0.117	0.139	0.334	0.132	0.125	0.020
24	8	0.111		0.117	0.144	0.335	0.140	0.129	-0.026
36	8	0.121		0.121	0.147	0.363	0.157	0.133	-0.057
48	8	0.122		0.120	0.146	0.350	0.148	0.125	-0.082
0	10	0.102		0.076	0.083	0.300	0.123	0.087	0.139
12	10	0.107		0.077	0.083	0.350	0.126	0.085	0.143
24	10	0.102		0.081	0.090	0.361	0.125	0.088	0.117
36	10	0.092		0.077	0.085	0.346	0.113	0.090	0.090
48	10	0.086		0.078	0.087	0.305	0.118	0.083	0.063
0	11	0.080		0.043	0.042	0.358	0.090	0.052	0.190
12	11	0.083		0.046	0.043	0.354	0.093	0.053	0.192
24	11	0.084		0.047	0.045	0.349	0.088	0.051	0.187
36	11	0.080		0.049	0.046	0.347	0.082	0.051	0.172
48	11	0.077		0.049	0.047	0.344	0.083	0.055	0.162

We wished to see how sensitive the regional predicted thunderstorm probability scores were to the model used to generate them. In column B, we show the results of using the region 6 model on the region 5 forecast data (column B, region 5), and region 5 model on region 6 forecast data (column B, region 6). This pair of regions was chosen for this experiment because of the significance of the difference in skill shown in column A for the two regions (that is, when their respective model was used). The fact that the results change very little from their column A counterparts in either table indicates that the model used was less important than the data set it was applied to and verified against. Region 5 has better scores than region 6 regardless of which model was used to diagnose probability. This suggests that, at least for these two regions, the diagnosis models are essentially interchangeable.

Another test of the sensitivity of the predicted probability skill to the diagnosis model used is to use the generalized model (that is, the model developed using the data for all regions) in all regions. The results for this experiment are given in column C. The degree of change of skill seems to vary by region. Only in region 4 for all thunderstorms does the skill level drop significantly when the generalized model is used instead of the regional model. In some regions (7 and 10, possibly 8 and 11 for all thunderstorms; 4,5,6,10, and 11 for severe thunderstorms) the skill level actually seems to improve. In the remaining regions there is no appreciable change in skill. Overall, this suggests that the skill obtained by the generalized model is no worse than, and possibly even better than, that from the application of regional models to their own regions. Here again, the model used seems to be less important than the data it is applied to in determining predicted probability skill.

In order to show a measure of relative skill of a forecast procedure, it is of interest to examine the skill obtained as compared with the skill resulting from a random forecast. This comparison

can be made by comparing the scores in columns A (C) with the scores in E when evaluating the skill of the regional (generalized) model. It is clear from the results that the model does show skill with respect to a random forecast of thunderstorm probability in both non-severe and severe scenarios. However, this differential is decreased when the random probabilities are chosen randomly from the pool of probabilities diagnosed by the regional (generalized) model, column F (column G), rather than purely randomly assigned (column E). The fact that the column F (G) scores are better than column E but worse than column A (C) suggests that: use of the model can narrow down the probability range from that of a random guess; there is even some skill associated with the location and time of the diagnosis as opposed to randomly choosing a probability that is typical of the continental US in August; a forecast from a typical probability range for a region and season (like that which could be derived just from the thunderstorm occurrence data) is probably better than a fully random guess. In a few of the regions, the skill from randomly selecting diagnosed probabilities from the generalized model (column G) is better than from randomly selecting probabilities diagnosed from the model applied to its respective region. In no cases is the opposite true. This seems to reinforce the suggestion that the generalized model is a suitable replacement for the regional model in essentially any region.

Column H gives the bias of the probabilities diagnosed by the regional models (that is, remove the squared factor in the above equation). In the all-thunderstorms validation, the results show that there is a tendency to underpredict the probability of thunderstorms in regions 7, 8, and 10. There seems to be no tendency to overpredict non-severe thunderstorms in any of the regions. Just the opposite trend is observed for severe thunderstorms. The tendency to overpredict occurrence of severe thunderstorms is expected since their actual frequency of occurrence is so low.

In summary, the validation results show that the skill of predicting the probability of thunderstorm occurrence by the perfect prog models developed in this project is relatively insensitive to the model used. The difference in skill between regions is greater than the difference in skill when two different models are applied to the same region. The results suggest that the generalized model can be substituted for a region's model without any appreciable loss of skill (and perhaps even an improvement in skill) in that region. The diagnosis models show skill with respect to purely random probability forecasts and probability forecasts assigned randomly from a range of reasonable probabilities. The tendency of the regional models to underpredict probabilities of all severities of thunderstorms is regionally dependent, but probably not model dependent. The models don't seem to have a tendency to overpredict except where the actual occurrence of thunderstorms is rare.

References

1. Reap, R.M., 1989: 24-H NGM Based Probability and Categorical Forecasts of Thunderstorms and Severe Local Storms for the Contiguous U.S., *NWS Technical Procedures Bulletin*, Series No. 419. 2.

2. Albers, S.C., 1987: A Severe Weather Forecast Software Package, *Proc. Symp. Mesoscale Analysis and Forecasting*, Vancouver, Canada, 17-19 August 1997.

3. Knapp, D., 1995, Personal Communication.

4. Hart, J.A., and J. Korotsky, *The SHARP Workstation v1.50 User's Manual*, NWS Forecast Office, Charleston WV.

5. McGinley, J., K. Runk, and J. Alleca, 1987: Use of Time Changes of Surface Data Parameters in Short Range Forecasting, *Proc. Symp. Mesoscale Analysis and Forecasting,* Vancouver, Canada, 17-19 August 1997.

6. Charba, J.P., 1979: Two- to Six-Hour Severe Local Storm Probabilities: An Operational Forecasting System, *Mon. Wea. Rev.*, **1007**:268-282.

7. Maddox, R.A., 1980: An Objective Technique for Separating Macroscale and Mesoscale Features in Meteorological Data, *Mon. Wea. Rev.*, **108**:1108-1121.

8. Storm Data and Unusual Weather Phenomena with Late Reports and Corrections, (1992 and 1993) National Climatic Data Center.